

Univerzita Karlova v Praze
Přírodovědecká fakulta
Katedra aplikované geoinformatiky a kartografie

Studijní program: Geografie
Studijní obor: Geografie a kartografie



Jan Šimbera

VYMEZOVÁNÍ FUNKČNÍCH REGIONŮ V PROSTŘEDÍ GIS

THE DELIMITATION OF GEOGRAPHICAL FUNCTIONAL REGIONS IN GIS

Bakalářská práce

Praha 2014

Vedoucí bakalářské práce: Mgr. Zuzana Žáková

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a že jsem všechny použité prameny řádně citoval. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Jsem si vědom toho, že případné použití výsledků, získaných v této práci, mimo Univerzitu Karlovu v Praze je možné pouze po písemném souhlasu této univerzity. Svoluji k zapůjčení této práce pro studijní účely a souhlasím s tím, aby byla řádně vedena v evidenci vypůjčovateli.

V Praze 19. května 2014

Jan Šimbera

Poděkování

Své vedoucí práce, Mgr. Zuzaně Žákové, děkuji za věnovaný čas, systematickou podporu a hodnotné připomínky pro tvorbu práce. Dále bych chtěl poděkovat za připomínky i vlastní impuls pro věnování se problematice RNDr. Miroslavu Maradovi, Ph.D., bez jehož projektu GA ČR č. P404-12-1035 (Prostorová dynamika dopravních vztahů v sídelním systému Česka) by tato práce nevznikla. Svě rodině bych pak rád poděkoval za materiální i morální podporu při tvorbě.

Abstrakt

Práce se zabývá řešením problému vymezení funkčních regionů, tedy regionů vymezených na základě interakcí (toků) v území, s použitím geografických informačních systémů (GIS). Jsou představeny základní přístupy popsané v české i světové literatuře, u nichž se práce se snaží poukázat na složitost jejich aplikace, zejména co se týče zahrnutí prostorových aspektů, a nalézt zjednodušení pomocí implementace v GIS. Je navržena vlastní metoda vymezení funkčních regionů, založená na přístupu M. Hampla; tato metoda je implementována spolu s dalšími pomocnými nástroji v prostředí ArcPy jako ucelený toolbox pro ArcGIS 10. Výsledky metody a implementace jsou kriticky zhodnoceny.

Klíčová slova: Funkční regiony, komplexní sociogeografické regiony, nodální regiony, problém vymezení funkčních regionů, regionalizace, GIS, ArcPy, sousedství

Abstract

The thesis deals with the functional areas delineation problem, seeking the best way to delimit functional regions based on spatial interactions (flows) using geographic information systems. The basic approaches from Czech and international literature are described. This thesis aims to show both their complicatedness and incompleteness, mainly concerning their spatial aspects, and to find a simplification by GIS implementation. A new method for delimiting functional regions is provided, based on the approach by M. Hampl. This method is implemented together with other auxiliary tools in the ArcPy environment as a toolbox for ArcGIS 10. The results of the method and its implementation are critically evaluated.

Keywords: Functional regions, functional areas delineation problem, local labour markets, nodal regions, regionalization, GIS, ArcPy, neighbourhood

Obsah

1 Úvod	8
2 Problém vymezení funkčních regionů	9
2.1 Koncept funkčních regionů	9
2.2 Formulace problému	10
2.3 Principy vymezení	12
2.4 Klasifikace vymezovacích metod	13
2.4.1 Kritéria vymezení	14
2.5 Prostorová dimenze vymezování	15
2.5.1 Sousedství	16
2.5.2 Vzdálenost	18
2.6 Existující metody vymezování	18
2.6.1 Hamplova regionalizace Česka	18
2.6.2 Intramax	20
2.6.3 Regionalizace CURDS	20
2.6.4 Bezákova regionalizace Slovenska	22
2.6.5 Metoda funkční vzdálenosti	22
2.6.6 Evoluční metody	23
2.6.7 Shrnutí	24
3 Návrh metody vymezení	25
3.1 Selekční kritérium	25
3.2 Agregáčn� krit�rium	25
3.2.1 Zahrnutí vzdálenosti	27
3.3 Ov�řovací krit�rium	28
3.3.1 �lenské funkce	28
3.3.2 �lenská v�ha	29
3.3.3 Stanoven� hranic	29
3.4 P�řezovací krit�rium	30
3.5 Slu�ovací krit�rium	30
3.6 Sousedství	31
3.7 Parametry metody	31
3.8 Vizualizace v�sledk�	32
4 Aplikace	33
4.1 Struktura a prvky toolboxu	33
4.2 Implementace vymezovac� metody	35

4.2.1	Struktura tříd	35
4.2.2	Vstupní data a parametry	35
4.2.3	Zajištění vstupu a výstupu	37
4.2.4	Implementace kritérií vymezení	38
4.2.5	Běh nástroje	40
4.3	Implementace ostatních prvků	41
4.3.1	Sousedství	41
4.3.2	Vzdálenost	43
4.3.3	Vizualizace příslušnosti zón k regionům	44
4.4	Postup použití	44
4.5	Regionalizace Česka	46
4.5.1	Použitá data a parametry	46
4.5.2	Výsledek	47
5	Diskuze výsledků	50
6	Závěr	52
	Použité zdroje	53
	Přílohy	57

Seznam tabulek

1	Tabulka interakcí použitá při vymezování – popis polí	46
2	Porovnání charakteristik vymezených regionů	47

Seznam obrázků

1	Příklad funkčního regionu	10
2	Problém MAUP při nevhodně zvolených zónách	12
3	Problém administrativního sousedství	17
4	Thiessenovy polygony zón	17
5	Hamplův koncept regionu	19
6	Hamplovo odvození generalizovaného modu	19
7	Coombesovo ověřovací kritérium	21
8	Vývojový diagram navrhovaného algoritmu	26
9	Struktura vytvořeného toolboxu	33
10	Dopravní sousedství – problém přechodu zevnitř	42
11	Dopravní sousedství – ukázka výsledků	43
12	Schéma postupu použití toolboxu	45
13	Vymezení funkčních regionů – porovnání s administrativním členěním	48
14	Vymezení funkčních regionů – porovnání s vymezením Hampl a Marada (2014)	48

1 Úvod

Funkční regiony jsou jedním ze základních konceptů sociální geografie. Jde o regiony vymezené na základě funkčních vztahů mezi jednotlivými částmi prostoru, tedy dynamických procesů (např. pohybů obyvatel), čímž se liší od regionů homogenních, vymezovaných staticky (např. dle převládajícího druhu průmyslové výroby). Snaží se v maximální míře odrážet přirozenou strukturu reality z hlediska člověka (spádovost), proto se také ve vzrůstající míře používají jako podkladové členění jak ve vědeckých studiích, tak ve státní správě (např. jako podklad pro návrh administrativního členění nebo alokaci zdrojů dekoncentrovaným pracovištím s určitou územní působností). Jak podotýkají Coombes a kol. (1978), použité územní členění může značně ovlivnit výsledek práce, a proto je nutné vybrat to odpovídající.

Tato práce se zabývá vymezováním těchto funkčních regionů se zvláštním důrazem na regiony sociogeografické, řeší tedy problém vymezení funkčních regionů (functional area delineation problem – FADP). V sociální geografii je tato problematika poměrně široce rozvinuta a zpracována v rozsáhlé literatuře; v Česku zaujímá dominantní postavení Hampl (2005), z jehož přístupu vychází i tato práce.

Naopak potenciál geoinformatiky zůstává v této oblasti nevyužit. Současné metody jsou realizovány ručně nebo s pomocí jednoduchých jednoúčelových programů, aniž by byla hlouběji využita prostorová dimenze vymezování, kterou lze jednoduše podchytit v prostředí geografických informačních systémů (GIS). Ty mohou být přínosné v několika oblastech:

- usnadnění práce s velkými objemy podkladových dat, které jsou pro zachycení prostorových vztahů nezbytné,
- umožnění integrace prostorových vztahů (sousedství, vzdálenost) do procesu vymezování, čímž odpadá nutnost přechodů mezi různými aplikacemi,
- snadnou vizualizaci výsledku, která umožňuje rychlou zpětnou vazbu zpracovateli.

Struktura práce je následující: V kapitole 2 je analyzována problematika vymezování sociogeografických funkčních regionů z geoinformatického pohledu a jsou diskutovány možnosti využití GIS v tomto procesu.

V kapitole 3 je na základě získaných poznatků navržena vlastní metoda vymezování funkčních regionů, která je následně v kapitole 4 implementována v prostředí ArcPy a v této podobě na ukázkou použita pro tvorbu funkční sociogeografické regionalizace Česka. V kapitole 5 jsou výsledky kriticky zhodnoceny.

2 Problém vymezení funkčních regionů

Funkční region je „územní celek vznikající na základě organizace vztahů v území, který je charakteristický vysokou intenzitou vnitřních prostorových interakcí“ (Konjar a kol. 2010, s. 1). Definicí založenou na vztazích se liší od regionů *homogenních*, definovaných jednotlívým statickým znakem, typicky hodnotami určitých veličin. Příkladem homogenních regionů mohou být klimatické oblasti, vymezované dle teploty a srážkových úhrnů, naproti tomu funkčními regiony jsou například regiony migrační, vymezované na základě vysokých intenzit vnitřní migrace.

S funkčními regiony pracuje především sociální geografie, na rozdíl od fyzické geografie, která vymezuje především regiony homogenní¹, a proto se tato práce věnuje *sociogeografickým funkčním regionům*², které jsou vymezovány především na základě interakcí mezi sídly. Hampl (2005) je nazývá *komplexní sociogeografické regiony*, Bezák (2000) *funkční městské regiony (FMR)*, v zahraničí se různě používají i termíny *travel-to-work areas* (TTWA – oblasti dojížděky za prací; Coombes a kol. 1986) nebo *local labour market areas* (LLA, LLM nebo LLMA – lokální trhy práce; van der Laan a Schalke 2001). Tato práce se primárně zabývá vymezováním těchto funkčních městských regionů. Funkční *regionalizací* se potom rozumí jednak proces vymezování funkčních regionů, tak i jeho výsledek – regionální členění (Hampl 2005).

Tato kapitola se věnuje nástinu regionalizačních metod používaných v českém i světovém kontextu s důrazem na prvky využitelné při implementaci řešení v GIS. Nejdříve je však nutné představit vlastní koncept funkčního regionu.

2.1 Koncept funkčních regionů

V současné literatuře (Sýkora a Mulíček 2009, Coombes a kol. 1986, Bezák 2000) existuje několik definic funkčního regionu, všechny však spatřují základní mechanismus jeho fungování v jeho *nodální* formě, která spočívá v polaritě dvou základních složek:

- *Jádro* (centrum, *core*), představující koncentraci aktivit, místo s vysokou atraktivitou. V případě funkčního městského regionu je představované jedním městem nebo několika vzájemně kooperujícími městy, z nichž je většinou jedno dominantní (Sýkora a Mulíček 2009).
- *Zázemí* (*hinterland*), které je s jádrem spojené sítí vazeb různého druhu, jejichž intenzita klesá se vzdáleností od jádra (Bezák 2000). Sýkora a Mulíček (2009) zdůrazňují dělení zázemí na dvě části – *funkční urbánní oblasti* (Functional Urban Areas – FUA, *Rings* – Coombes a kol. 1982), urbanizované oblasti těsně vázané na jádro, na němž

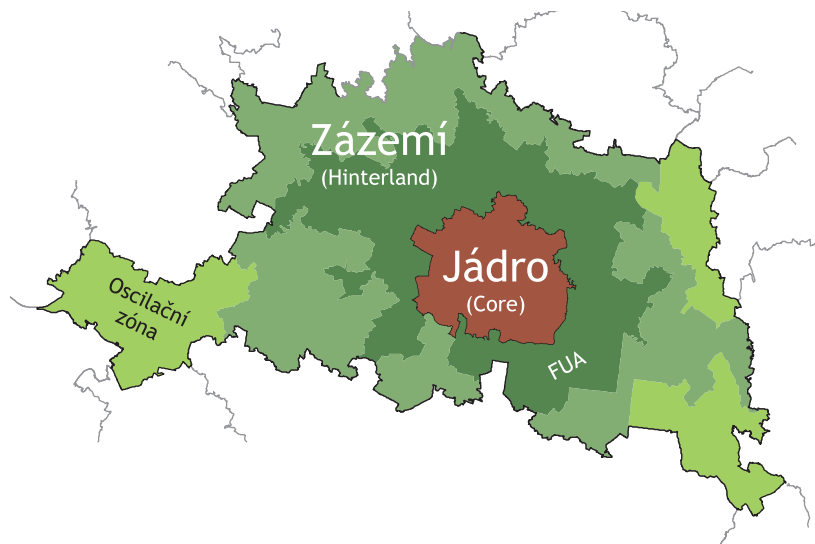
¹Existují samozřejmě výjimky obojího druhu – fyzickogeografické funkční regiony (např. povodí) nebo sociogeografické homogenní regiony (např. regiony převládajících ekonomických aktivit).

²Zde představené metody ovšem nevyklučují aplikaci na vhodná fyzickogeografická data.

jsou závislé ve většině funkcí, a slaběji integrované periferní oblasti (*Outer Areas, Rural Areas*), přecházející postupně až do *oscilačních zón* s nejasnou regionální příslušností.

Díky této polaritě, kde jádro funguje jako důležitý uzel, se funkční regiony často nazývají *nodálními*. Pokud je jeho jádro vícečlenné, lze region označit jako *multinodální*. Příklad funkčního regionu představuje obrázek 1.

Obrázek 1: Příklad funkčního regionu: dojíždkový mikroregion Hradce Králové. Intenzita interakcí postupně klesá se vzdáleností od jádra – nejvyšší je v oblasti FUA, nejnižší v oscilačních zónách.



Zdroj: ArcČR (2013), vlastní zpracování

2.2 Formulace problému

Je dána množina *zón* Z , $Z = \{z_i\}$, $|Z| = n$ je počet zón, a čtvercová matice jejich vzájemných *interakcí* $F = (f_{ij})_{i,j \in \hat{n}}$, kde f_{ij} je tok ze zóny i do zóny j . Cílem je vytvořit množinu regionů $M = \{R_i\}$ ($|M| = k$ předem neznámé) tak, že $\forall i; R_i \subset Z, R_i \neq \emptyset, \bigcup_{i=1}^k R_i = Z$.

Zóny představují základní prostorové jednotky, za něž jsou dostupná příslušná interakční data. Jde o obecný technický pojem, jež používá Bezák (2000) pro elementární prvky analýzy, aby bylo možné abstrahovat od konkrétního národního sídelního systému; nejde o oblasti vymezené na základě homogenity určitého jevu. Například v Česku jsou jako zóny používány obce, o nichž jsou dostupná dostatečná data (prostorové vymezení, počet obyvatel, interakce).

Každé zóně je přiřazena určitá váha m_i (*mass*) – nejčastěji jde o populační velikost (veškeré nebo pouze ekonomicky aktivní populace). Prostorové určení zón vstupuje do algoritmu nejčastěji jako matice sousedství $C = (c_{ij})_{i,j \in \hat{n}}$, kde $c_{ij} = 1$, pokud spolu zóny i a j sousedí, jinak 0.

Guoqing (2001) uvádí jako hlavní *interakce* mezi sídly toky obyvatel, zboží, kapitálu a dat. V reálné situaci je právě dostupnost dat o těchto interakcích zásadním omezením funkční regionalizace, neboť je značně obtížné tato data získat. Pohyb obyvatelstva v území není systematicky sledován a podobně tomu je i u jiných prostorových aktivit (např. telefonních hovorů); takovéto sledování nutně naráží na ochranu soukromí, osobních údajů

nebo obchodního tajemství. Podrobnější data lze získat v rámci případových studií od dobrovolně spolupracujících jedinců – u nás tento postup použili např. Kunc a kol. (2013) pro vymezování regionů dle dojížděky za službami. V zemích, kde se dojížděka nesčítá plošně, je dotazníkové šetření jediným možným východiskem (viz např. Karlsson a Olsson 2006), ve větším měřítku je tento postup ovšem finančně neproveditelný.

Výzkumník tak zůstává omezen prakticky na jediné plošné dotazníkové šetření – sčítání lidu, kde se shromažďují údaje o směrově určené dojížděce za prací a do škol. Tato data však mají několik dalších nevýhod. Nejsou mnohdy v dostatečném územním rozlišení, protože obce jsou často značně heterogenní jednotky, skládající se z více sídel. Jsou také zatížena chybami, jimiž trpí každé dotazníkové šetření, tím více to, které není založeno na dobrovolnosti. Jistou alternativu nabízí sčítání intenzit silniční dopravy (u nás prováděné ŘSD), u nichž vyvstává problém s převodem na interakce, analýza spojení veřejné hromadné dopravy, která je však do značné míry ovlivněna administrativním členěním a neposkytuje úplný obrázek, nebo statistiky migrace, použitelné pouze na vyšších řádovostních úrovních.

Pro vymezování funkčních městských regionů nejnižší (mikroregionální) úrovně se v drtivé většině případů používá denní dojížděka za prací, jež se v rámci sčítání lidu běžně zjišťuje. Na tomto principu je postaven i koncept *denního urbánního systému* (DUS), z něž čerpá většina zahraničních prací. Panuje přitom všeobecná shoda, že pracovní dojížděka je jedním z nejvýznamnějších pohybů obyvatel a navíc se jí přizpůsobují i pohyby ostatní, např. za službami (Bezák 2000); má však i svá specifika (někteří lidé dojíždí za prací velmi daleko, takže jejich pohyb neodráží regionální spádovost). Tato práce pracuje primárně právě s dojížděkovými interakcemi³.

Problém nastává v okamžiku, kdy jsou interakční data dostupná pouze za jednotky vyššího řádu než sídla (administrativní jednotky). Výsledná regionalizace pak trpí problémem MAUP⁴, které nelze rozumně eliminovat (viz obr. 2). V Česku nastává takový problém naštěstí pouze marginálně u obcí skládajících se z většího počtu funkčně nezávislých sídel (prostorová nespojitost, odlišná spádovost); ve většině zahraničních studií však přílišná velikost základních jednotek výrazně zkresluje výsledky (Karlsson a Olsson 2006). Naopak data v jemnějším detailu než nejnižší úroveň administrativního členění jsou dostupná pouze výjimečně (v Anglii a Švédsku – viz např. Landré 2012).

Matice interakcí je většinou nesymetrická (záleží na povaze interakcí; některé interakce, např. telefonní hovory, jsou ze své podstaty symetrické, některé metody také pracují s obousměrnými interakcemi). Z hlediska zpracování je právě velikost této matice překážkou pro intuitivní analýzu; objem dat v ní roste kvadraticky s počtem zón⁵ ($\mathcal{O}(n^2)$). Počítačové zpracování se proto přímo nabízí.

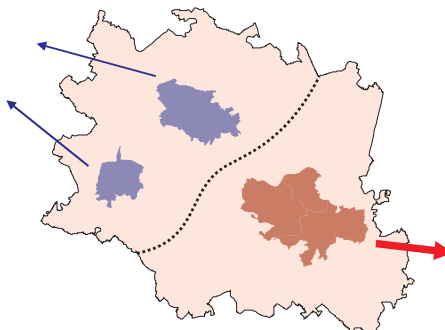
Označme $i_i = \sum_{k=1}^n f_{ki}$ součet hodnot všech interakcí končících v zóně i (příchozí proudy, *inflows*) a $o_i = \sum_{k=1}^n f_{ik}$ součet hodnot všech interakcí v zóně i počínajících (odchozí proudy, *outflows*). Interakce jednotlivé zóny lze také vyjádřit interakčním vektorem \vec{f}_i , představujícím řádek i interakční matice.

³Vedle dojížděky za prací se v Česku sleduje rovněž dojížděka do škol, která vykazuje poněkud jiný prostorový vzorec a je značně ovlivněná administrativně, proto je její význam pouze částečný (Hampel 2005); Hampel a Marada (2014) ji však s pracovní dojížděkou slučují kvůli nekompletnosti dat.

⁴Modifiable Area Unit Problem – problém vymezení odpovídajících územních jednotek. Používá se jednak pro skupinu problémů věnujících se vymezování územních jednotek, jednak pro obtíže při analýzách používajících nevhodné územní jednotky.

⁵Byť je v důsledku prvního zákona geografie (Tobler 2004) většinou řídká (podobně jako matice sousedství) a zdaleka všechny dvojice zón nejsou propojeny nenulovou interakcí.

Obrázek 2: Problém MAUP při nevhodně zvolených zónách. Zóna na obrázku obsahuje tři sídla, jejichž převládající směr interakcí je odlišný; interakce největšího sídla (červeně) při posuzování zóny jako celku převládá a tečkovaná hranice, jež by oddělovala regiony při použití podrobnějších dat, se posune vlevo.



Zdroj: ArcČR (2013), vlastní zpracování

Cörvers a kol. (2009) charakterizují problém vymezení funkčních regionů jako typ problému MAUP, skládajícího se z určení řádovosti (*scale problem* ústící v požadovaný počet regionů) a regionalizační metody (*aggregation problem*). Fortunato (2010) k jeho řešení poznamenává, že počet možných variant vymezení regionů roste exponenciálně s počtem zón, a proto není pro obvyklou velikost vstupu možné ověřit všechny varianty; místo toho je podle něj nutné uchýlit se k heuristikám. GIS takové heuristiky díky zahrnutí prostorové dimenze umožňuje sestavit kvalitnější a efektivnější, a tak vymezovat funkční regiony lépe odpovídající reálným prostorovým vztahům.

2.3 Principy vymezení

Hampl (2005) uvádí tři základní pojetí regionalizace.

- *Metodologické*, kde je region brán jako účelová jednotka pro další analýzy. Kritéria vymezení zde stanovuje analytik.
- *Realistické*, snažící se zjistit fakta o realitě jako takové na základě funkčních vztahů. Na jeho základě by všichni teoreticky měli dojít ke stejnému výsledku, odrážejícím reálný stav; v praxi je toto limitováno naší schopností vytvořit model reality, z nějž se odvozují kritéria vymezování.
- *Pragmatické*, vymezující regiony za konkrétním praktickým účelem (administrativní a plánovací regiony, volební obvody). K modelu reality zde mohou být dodána další kritéria vyplývající z požadavků na výsledek (např. snaha o územní spravedlnost, projevující se v požadavku územní celistvosti a srovnatelné velikosti regionů).

Sám se věnuje v maximální míře realistickému pojetí. V zahraničních studiích není pojetí v tomto smyslu definováno; nejčastější je snaha vymezit prakticky použitelné regiony, které budou v maximální míře odrážet reálný stav – jde tedy o průnik druhého a třetího pojetí.

Bezák (2000) (i např. Casado-Díaz a Coombes 2011) v tomto duchu uvádí osm obecných *principů* pro vymezování funkčních městských regionů. Z nich jsou pro proces vymezování nejpodstatnější následující čtyři:

Princip jednoznačnosti. Každé území by mělo náležet *právě jednomu* regionu – funkční regiony mají vyplňovat celé zkoumané území ($\forall z \in Z; \exists R \in M; z \in R$) a přitom se nepřekrývat ($\forall R_i, R_j \in M; i \neq j \Rightarrow R_i \cap R_j = \emptyset$). V případě opaku by obojí působilo problémy při následném používání takové regionalizace jako výchozího členění jak pro analýzy, tak pro administrativu. Realita je však spojitá a oblasti vlivu sousedních jader regionů se téměř vždy překrývají. Někteří autoři (van der Laan a Schalke 2001, Farmer a kol. 2011) také poukazují na rozdílnost regionů pro různé sociální skupiny. Toto překrývání souvisí s hierarchickým uspořádáním sídelního systému – často se proto tvoří regionalizace víceúrovňové (Hampl 2005, Coombes a kol. 1986), pokrývající více řádovostních úrovní.

Přímo proti tomuto principu jde teorie fuzzy regionů (Feng 2009, Watts 2013), která pracuje na principu fuzzy množin, kde prvek může patřit do množiny jen do určité míry, a přiřazuje každou zónu do více regionů v závislosti na jejich funkční provázanosti s daným regionem. V žádné z citovaných prací však není tato teorie použita pro vymezení regionů, pouze pro hodnocení existující regionalizace. Fuzzy regiony jsou však pro svůj charakter odporující výše uvedeným podmínkám využitelné pouze při realistickém pojetí regionalizace, a to ještě obtížně.

Princip souvislosti. Každý region by měl být prostorově souvislý bez enkláv a exkláv⁶ (jde o častý požadavek vycházející z pragmatického pojetí). Tento princip je z velké míry zajištěn prostorovou povahou samotných interakcí, jejichž síla se vzdáleností klesá, a tak jsou výsledkem relativně souvislé regiony i v případě, že v postupu nebyla podmínka souvislosti explicitně vyjádřena. Zejména v oblastech s dynamickým vývojem a velkou hustotou interakcí však dochází ke vzniku exkláv. Ty je nejvýhodnější eliminovat až v závěru; podmínka souvislosti při samotném vymezení může vést k problematickým výsledkům (Coombes a kol. 1986, Bezák 2000).

Princip autonomie. Region by měl tvořit autonomní jednotku systému co do interakcí zón v něm obsažených – maximum toků by mělo probíhat v rámci regionu (*vnitřní soudržnost*) a minimum mezi regionem a okolím (*vnější uzavřenost*). Tento princip tvoří jádro postupu vymezení. Vnitřní soudržnost se dosahuje agregací zón s maximální intenzitou interakce, jde tedy o maximalizaci určitého *agregačního kritéria* v průběhu přiřazování zón k regionům. Naopak zajistit vnější uzavřenost je větší problém, související s principem jednoty řádovosti.

Princip jednoty řádovosti. Všechny regiony, které jsou zároveň vymežovány, by měly být velikostně porovnatelné, tedy odpovídat stejné řádovostní úrovni (přičemž jedno jádro může mít na různých řádovostních úrovních různě rozsáhlé regiony). Toho se zpravidla dosahuje stanovením určitých minimálních požadavků na daný region (minimálního počtu obyvatel, vnější uzavřenosti) – tedy určitého *ověřovacího kritéria*.

2.4 Klasifikace vymežovacích metod

O klasifikaci regionalizačních metod se pokoušeli van der Laan a Schalke (2001) a nověji Casado-Díaz a Coombes (2011). Jedno z dělení je založeno na rozdílu metod

⁶Exkláva je zóna nebo skupina zón, která nesousedí se zbytkem svého regionu, kde se nachází regionální jádro.

- *deduktivních*, používajících selekční kritérium (viz 2.4.1) a zdůrazňujících nodální charakter regionu (Hampl 2005, Konjar a kol. 2010), a
- *induktivních*, které selekční kritérium nepoužívají a jádra vymezují pouze na schopnosti přitáhnout dostatečné zázemí (většina moderních metod).

Z hlediska vlastního běhu algoritmu je významné dělení na metody

- *hierarchické*, používající po celou dobu běhu pouze jedno kritérium (*agregační*, viz 2.4.1), jako např. Intramax (Masser a Brown 1975) nebo top-down partitioning (Farmer a kol. 2011), a
- *založené na pravidlech (rule-based)*, které používají kritérií více; zejména stanovují jasná *ověřovací kritéria* pro výsledné regiony. Mezinárodně nejznámější vyvinuli Coombes a kol. (1986), lze sem však řadit i většinu českých (Hampl a Marada 2014, Sýkora a Mulíček 2009).

Hierarchické metody vycházejí spíše z teoretických předpokladů a stojí na matematických základech, čímž se přibližují metodám detekce komunit (*community structure detection*) z teorie grafů (více viz Fortunato 2010, pro nejčastější spektrální dělicí metody – *spectral partitioning* – viz Nascimento a kol. 2011). Jejich výhodou je jasné odůvodnění všech kroků a většinou i jednoduchost. Naopak jsou často kritizovány, že neodráží empirické zákonitosti tvorby vztahů v území (Findlay a Slater 1981) a jejich jednostupňový charakter nedostatečně podchycuje jejich změny na různých řádovostních úrovních, když nedokáže rozloučit jednou spojené zóny (Casado-Díaz a Coombes 2011).

2.4.1 Kritéria vymezení

Každá regionalizační metoda pracuje na základě určitých kritérií⁷. Každé z nich podléhá do určité míry subjektivní volbě tvůrce metody – často jsou tvořena teprve testováním jejich výsledků na zpracovávaných datech (Bezák 2000). V tom spočívá hlavní příčina nekompatibility regionalizací v různých státech. Cörvers a kol. (2009) a Farmer a kol. (2011) proto nedoporučují používat jakékoli explicitní číselné míry.

Tato práce se snaží kritéria používaná v různých pracích klasifikovat do pěti druhů.

Selekční kritérium vybírá z množiny zón množinu jader C ($C \subset Z$). Aplikuje se u těch metod, které vyžadují na začátku běhu explicitní rozdělení zón na jádra a zázemí, a má nejčastěji formu minimální populační velikosti nebo lokalizačního kvocientu pracovních míst:

$$LQ_i = \frac{w_i \sum_{k=1}^n m_i}{m_i \sum_{k=1}^n w_i},$$

kde w_i je počet zaměstnaných v i (někdy omezených sektorem) a m_i váha této zóny. Cílem je hned ze začátku neuvažovat příliš malé, nevýznamné nebo závislé zóny bez dostatečného počtu pracovních příležitostí, a tak celý postup vymezení zjednodušit a zrychlit.

Casado-Díaz a Coombes (2011) i další selekční kritérium kritizují pro jeho apriorní charakter a model osídlení, který „je dnes odsunut do pozadí růstem vícejaderných sídelních

⁷Ty je nutno odlišit od principů; zatímco principy jsou obecné, kritéria jsou exaktní a zpravidla jednoznačně kvantifikovatelná.

systémů a decentralizací nabídky zaměstnání“ (Casado-Díaz a Coombes 2011, s. 13). Také tvrdí, že pro počítačové zpracování není v současnosti nutné díky dostatečnému výkonu. Selekční kritérium lze také odmítnout na základě myšlenky, že schopnost zóny tvořit jádro regionu lze nejlépe hodnotit právě její schopností přitáhnout a udržet dostatečné zázemí.

Agregační kritérium se vytváří nad silou interakce a určuje pravděpodobnost, s níž budou dvě zóny přiřazeny stejnému regionu. Zpravidla zahrnuje kromě síly interakce i charakteristiky počáteční a koncové zóny ($o_i, i_j; m_i, m_j$). Toto kritérium je v průběhu metody maximalizováno – zóna je přiřazena k tomu regionu, pro nějž je hodnota jejího agregačního kritéria nejvyšší.

Nejjednodušším kritériem je sama velikost interakce (Hampl 2005) (*metoda primárního toku*); van der Laan a Schalke (2001) relativizují tuto velikost dělením nejsilnější interakcí zóny.

Farmer a kol. (2011) navrhuje pro zohlednění principu jednoty řádovosti aplikovat na agregační kritérium korekci dle vzdálenosti (*distance-decay*), přes kterou interakce probíhá, aby bylo možné odfiltrovat interakce jiných řádovostních úrovní než té, na níž je daná regionalizace konstruována.

Ověřovací kritérium popisuje vytvářený region a určuje jeho kvalitu. Pokud region kritérium nesplní, je podle pravidel metody rozpuštěn a jeho zóny přiřazeny jinam. Ověřovací kritérium obsahují všechny metody založené na pravidlech. Nejčastěji je jejich obsahem velikostní význam (Hampl (2005) používá odděleně i velikost zázemí) a vnější uzavřenost, které se často mohou do určité míry vzájemně kompenzovat (*trade-off*), takže větším regionům je tolerována nižší uzavřenost.

Přeřazovací kritérium je aplikováno pro optimalizaci výsledného řešení. Jeho cílem je maximalizace ověřovacího kritéria tak, že přeřazuje zóny (většinou pouze hraniční) z jednoho regionu do druhého. Postup je analogický relokačním procedurám používaným ve shlukové analýze (Bezák 2000).

Slučovací kritérium slučuje dohromady dva funkčně silně propojené regiony na základě intenzity vzájemné interakce jejich jader nebo celých regionů. Vzniká tak region multino-dální.

2.5 Prostorová dimenze vymezení

Přestože síla interakcí v souladu s prvním zákonem geografie se vzdáleností přirozeně klesá (Tobler 2004), metody nezahrnující žádným způsobem prostorovou dimenzi často vytvářejí regiony nesouvislé nebo přiřazují některé zóny příliš vzdáleným střediskům. Z toho lze usuzovat, že tento přirozený pokles síly se vzdáleností není rovnoměrný.

Nevhodnost nesouvislých regionů přímo vychází z principu souvislosti; nevhodnost přiřazení zóny příliš vzdálenému jádru pak vyplývá implicitně z požadavku vnitřní soudržnosti principu autonomie (příliš rozsáhlé regiony tuto soudržnost nutně ztrácejí) a z principu řádovostní srovnatelnosti. Interakce ve vstupních souborech totiž často odpovídají více řádovostním úrovním (například dojíždka za vysokými školami), a je tudíž třeba odfiltrovat

ty, které svou řádovostní úroveň neodpovídají vymezeným regionům. Výsledné regiony pak budou spíše velikostně odpovídat povaze interakcí, které je na dané řádovostní úrovni integrují.

Pokud má být princip souvislosti aplikován důsledně, je nutno prostorovou dimenzi do metody zahrnout, byť jsou výsledná řešení z hlediska ostatních kritérií suboptimální – jinak by se pochopitelně vytvořila i bez zahrnutí prostorové dimenze (Casado-Díaz a Coombes 2011). Lze přitom využít dvou konceptů: *sousedství* a *vzdálenosti*.

GIS je pro aplikaci prostorové dimenze ideálním prostředím, neboť přímo implicitně pracuje s geometrií vstupních dat. Výhoda integrace celého vymezení procesu do GIS netkví primárně v rozšíření teoretických možností – prostorové vztahy sousedství a vzdálenosti lze samozřejmě ve vhodném datovém formátu z GIS exportovat a použít v jiném prostředí – ale v jednoduchosti a přímočarosti integrace všech potřebných kroků od zpracování dat až po vizualizaci výsledku v jednom prostředí. Odpadají tak datové konverze, které jsou často nejdéletrvající částí celého procesu a také nejnáchylnější k chybám.

2.5.1 Sousedství

Většina existujících metod zahrnuje koncept sousedství jako lokální prostorový vztah blízkých zón pro zajištění souvislosti regionu. Buď je podmínka sousedství pevně součástí agregačního kritéria, nebo jsou exklávy eliminovány dodatečnou (často dosti subjektivní) úvahou. V žádné z nalezených metod se však nevyskytuje postupná aplikace této podmínky, kdy jsou nejprve regiony vymezeny bez ohledu na sousedství, načež je podmínka sousedství přidána a regionalizace se upraví způsobem analogickým k procesu vymezení (např. pomocí přerazovacího kritéria).

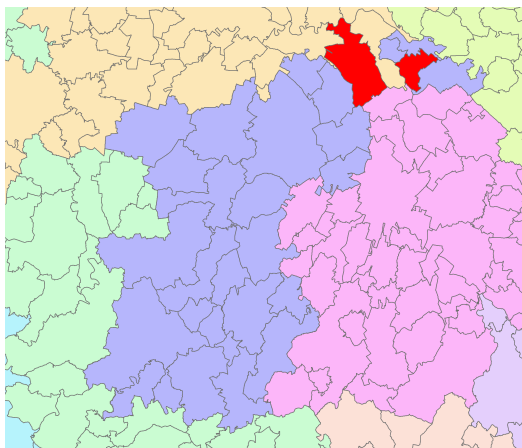
Existuje několik způsobů, jak definovat sousedství dvou zón:

Administrativní sousedství. Jde o nejjednodušší přístup, kdy sousedí zóny, jejichž polygony mají společnou hranici (pro větší robustnost se obvykle ověřuje, zda mají nenulový průnik). Používá se nejčastěji (Burrough a McDonnell 1998) – zejména proto, že prostorový tvar výsledného regionu je vždy souvislý, pokud se ovšem polygony neskládají z více částí. To je občas problémem u administrativních jednotek nižšího řádu zahrnujících více sídel (např. českých obcí, jichž je přibližně sto prostorově nesouvislých; viz obr. 3). Přirozeně tak vyvstává otázka, zda nedefinovat sousedství pomocí jiného průběhu hranic než administrativních.

Thiessenovy polygony (též Voronnoiovy buňky) se v tomto ohledu přímo nabízí použít. Každý bod prostoru je pomocí nich přiřazen té zóně, k jejímuž centrálnímu bodu má nejbližší (Worboys a Duckham 2004; pro ukázkou viz obr. 4). Centrální bod bývá umístěn v centru hlavního sídla tvořícího zónu, tedy často odlišně od centroidu polygonu administrativních hranic zóny. Pokud data o centrálních bodech zón nejsou k dispozici, lze centroidy použít – výhodou je blízkost výsledku administrativním hranicím, nevýhodou méně pravdivá informace o prostorových vztazích vlastních sídel zóny tvořících.

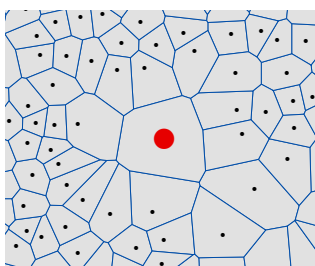
Thiessenovy polygony však trpí nevýhodou v oblastech kolem velkých zón. Ty jsou reprezentovány jediným centrálním bodem, který je v důsledku velikosti zóny značně vzdálen

Obrázek 3: *Problém při použití sousedství dle administrativních hranic zón. Region uprostřed je nesouvislý, přestože při vymezení byla striktně aplikována podmínka sousedství; je totiž nesouvislá zvýrazněná zóna.*



Zdroj: ArcČR (2013), vlastní zpracování

Obrázek 4: *Thiessenovy polygony centrálních bodů obcí v okolí Hradce Králové (uprostřed). Je zřejmé, že plocha polygonu Hradce Králové (a tedy i množství obcí, s nimiž sousedí) neodpovídá jeho významu ve srovnání s okolními vesnickými obcemi.*



Zdroj: ArcČR (2013), vlastní zpracování

od centrálních bodů zón, s nimiž sousedí administrativní hranicí. Řešením mohou být dva způsoby:

Vážené Thiessenovy polygony. Vzdálenost je při tomto přístupu dělena masou příslušné zóny, takže zóny s velkou masou mají širší sféru vlivu. Výsledkem jsou vážené Thiessenovy polygony (Dong 2008), které odpovídají aplikaci Reillyho gravitačního modelu. Váhu zóny je však nutno odpovídajícím způsobem upravit, aby odpovídala její regionální působnosti; lze při tom využít početnou literaturu o gravitačním modelování. Kraft a Blažek (2012) navrhuji na základě českého sídelního systému použít pátou odmocninu.

Výplň sekundárními body. Každá zóna je dle pravidelné mřížky o velikosti odpovídající průměrné velikosti zóny vyplněna body, nad nimiž se vytvoří Thiessenovy polygony. Ty jsou poté sloučeny na základě příslušnosti zdrojového bodu v zóně.

Sousedství založené na dopravních sítích je další alternativní způsob, autorem vyhodnocený jako nejpřiléhavější, neboť nejlépe odpovídá reálnému vnímání sídelního systému

z hlediska člověka. Dvě zóny spolu sousedí, pokud mezi jejich centrálními body existuje dopravní spojení neprocházející žádným sídlem, které nepatří do jedné z obou zón – pokud tedy musím při cestě ze zóny a do zóny b projet i sídlem ze zóny c , zóny a a b spolu nesousedí. Taková úloha patří svou povahou do síťové analýzy.

2.5.2 Vzdálenost

Koncept vzdálenosti jako prostředek zajištění vnitřní soudržnosti se v existujících metodách vyskytuje poměrně zřídka. Farmer a kol. (2011) na interakce variantně aplikuje gaussovskou vzdálenostní korekci ve formě

$$f'_{ij} = f_{ij} \cdot \exp\left(\frac{-d_{ij}^2}{b}\right),$$

kde f_{ij} je původní tok mezi zónami i a j , d_{ij} vzdálenost jejich centrálních bodů a b je parametr určující rychlost úbytku síly interakce se vzdáleností, který může být definován manuálně nebo určen na základě analýzy vstupních interakcí (což používá i Farmer). Tímto způsobem jsou penalizovány interakce působící na nadprůměrně velkou vzdálenost. Místo vzdušné vzdálenosti lze samozřejmě opět využít vzdálenost po dopravních sítích.

Další možností by bylo využít gravitačního modelu, který zahrnuje i masy interagujících zón, a oslabit interakce působící na příliš velkou vzdálenost vzhledem k masám jejich středisek.

2.6 Existující metody vymezení

V této části je popsáno několik regionalizačních metod, které jsou široce používané buď v Česku, nebo v zahraničí, případně jsou obzvláště pozoruhodné některým ze svých prvků.

2.6.1 Hamplova regionalizace Česka

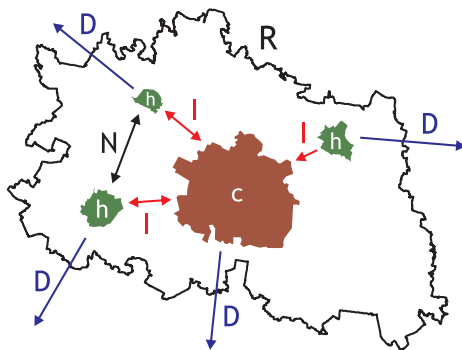
Sociogeografické regionalizace Česka z pera Martina Hampla (Hampl a kol. 1987, Hampl 1996, 2005, Hampl a Marada 2014) jsou nejkompletnějším zachycujícím český sídelní systém, navíc také pokrývají nejdelší časový úsek (již od roku 1960), čímž umožňují vývojová srovnání. Vlastní vymezení funkčních regionů slouží jen jako rámec pro hodnocení sídelního systému, v němž autor vidí hlavní hodnotu celé činnosti. Celá práce vychází z vlastní značně specifické teoretické báze (Hampl 1998), která nabízí několik zajímavých východisek, představených níže.

Regionalizace není vedena přesným algoritmem, nýbrž jen soustavou několika pravidel, doplněných autorovou důkladnou znalostí lokálních poměrů. Každá regionalizace přitom vychází z předchozích – v nejednoznačných případech je často použito minulých výsledků (stejně i v případě selekčního kritéria).

Hamplův koncept sociogeografického (funkčního) regionu zdůrazňuje polaritu jádra a zázemí a nepovažuje na rozdíl od zahraniční literatury toky v zázemí regionu za integrující, nýbrž za neutrální (viz obr. 5).

Selekční kritérium je aplikováno na základě výsledků minulých regionalizací, agregační kritérium je založeno na primárním toku, slučovací kritérium je použito v individuálních případech na základě subjektivní analýzy.

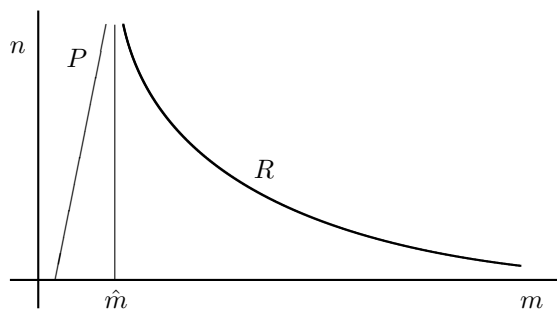
Obrázek 5: Hamplův koncept regionu. Toky označené I region R integrují, D dezintegrují, N jsou neutrální.



Zdroj: ArcČR (2013), vlastní zpracování

Ověřovací kritérium stojí na pevných empiricky stanovených hodnotách masy regionu a jeho zázemí (15 000, resp. 5 000)⁸. Hodnota 5 000 je považována za minimální pro formování polarity jádro-zázemí, 15 000 je určena jako *generalizovaný modus* velikostí regionů s takto velkým zázemím. Použití generalizovaného modu jakožto aproximace modu krajně asymetrického rozložení vychází z požadavku zařazení jádra do hierarchie ve smyslu Zipfova pravidla velikostního pořadí (Auerbach 1913, viz obr. 6). Pokud region ověřovací kritérium nesplní, je agregován podle *zásady skladebnosti*. Ta stanovuje, že region má být agregován jako celek k jinému regionu na základě primárního toku svého jádra.

Obrázek 6: Hamplův koncept hierarchie regionů pro odvození generalizovaného modu \hat{m} . R označuje běžnou hierarchii regionů dle Zipfova pravidla, P regiony s problematickou autonomií nezapojené do hierarchie.



Pro hodnocení uzavřenosti (integrity) regionů používá M. Hampl dvě míry – *integrity zázemí* I_z a *integrity regionu* I_r :

$$I_z(R) = \frac{\sum_{i \in h(R), j \in c(R)} f_{ij}}{\sum_{i \in h(R), j \notin R} f_{ij}}$$

$$I_r(R) = \frac{\sum_{i \in h(R), j \in c(R)} f_{ij} + \sum_{i \in c(R), j \in h(R)} f_{ij}}{\sum_{i \in R, j \notin R} f_{ij}},$$

kde $c(R)$ je jádro a $h(R)$ zázemí regionu. Tyto míry však nejsou součástí ověřovacího kritéria,

⁸Jedno z těchto kritérií může být sníženo na 10 000, resp. 2 500 – jde potom o nižší rádivostní úroveň, *subregion*.

pouze slouží k hodnocení výsledných regionů.

Hamplova regionalizace je unikátní svou propracovanou hierarchií regionů (v některých verzích až 7 stupňů). Tyto regiony jsou získány na základě zásady skladebnosti; zatímco na nižších úrovních představuje vstupní interakce dojíždka za prací, na vyšších úrovních se zohledňuje i dojíždka do škol a migrace. Taková konfigurace vybízí k několikastupňové aplikaci jednoho algoritmu s rozdílnými vstupními daty a hodnotami ověřovacího kritéria.

2.6.2 Intramax

Metodu Intramax vyvinuli Masser a Brown (1975). Je klasickým příkladem hierarchické agregační metody, která pracuje na principu postupného slučování zón.

Matice interakcí je nejprve standardizována dle principu rozdílu reálné hodnoty a hodnoty očekávané dle marginálních součtů, čímž se obdrží agregační kritérium:

$$s_{ij1} = \frac{f_{ij}}{o_i i_j} + \frac{f_{ji}}{o_j i_i}$$

kde f_{ij} je tok ze zóny i do zóny j , o_i součet odchozích toků ze zóny i a i_i součet příchozích toků do zóny i .

V každém kroku metody jsou sloučeny dvě zóny, jimž odpovídá aktuálně nejvyšší hodnota agregačního kritéria v matici, pokud splňují podmínku spojitosti. Není stanovena žádná zastavovací podmínka; metoda může běžet až do konečného vytvoření jediného regionu, proto je pro získání smysluplného členění nutné definovat žádaný počet regionů nebo podíl vnitroregionálních toků.

Metoda sbírá body díky své jednoduchosti a existující počítačové implementaci ve volně dostupném softwaru FlowMap (de Jong a kol. 2013). Její ukotvení v geografickém kontextu je velmi slabé, takže ji lze použít např. pro vymezení informačních sémantických vazeb (Fabrikant 2001). Dále tuto metodu v různých zemích využili např. Feldman a kol. (2005), Landré (2012), Landré a kol. (2013), Krygsman a kol. (2009), Mitchell a Watts (2010), Nielsen a kol. (2008), Watts (2013).

Kritika metody (Findlay a Slater 1981) je zaměřena na její hierarchickou povahu – především na nemožnost rozdělit jednou sloučené zóny (čímž často vznikají neorganické jednotky), nejasnost a malou flexibilitu zastavovací podmínky nebo přílišnou rigiditu podmínky spojitosti.

2.6.3 Regionalizace CURDS

Regionalizační metoda založená na pravidlech, za níž stojí Coombes a kol. (1986) z výzkumného centra CURDS (odtud metoda CURDS), kteří ji poprvé použili ve Velké Británii, získala velké mezinárodní uznání a mnoho autorů ji aplikovalo v jiných zemích (Andersen 2002, Bezák 2000, Casado-Díaz 2000, Newell a Perry 2004, Watts 2013), Česko nevyjímaje (Klapka a kol. 2013); často s různými modifikacemi. Zde je popsána její současná verze, oproti originálu zjednodušená (Bond a Coombes 2007).

Selekční kritérium není aplikováno (ve dřívějších metodách bylo obsaženo ve formě lokalizačních kvocientů kvůli urychlení). Agregační kritérium je založeno na široce rozšířené

míře podobné Intramaxu, již poprvé zavedl Smart (1974):

$$s_{ij2} = \frac{f_{ij}^2}{o_i i_j} + \frac{f_{ji}^2}{o_j i_i}$$

Tato míra zohledňuje vzájemný vztah zón vzhledem k jejich celkovému významu podobně jako Intramax.

Pro ověřovací kritérium definují Coombes a kol. (1986) dva typy uzavřenosti regionu: nabídkovou $C_S(R)$ (vzhledem k bydlišti) a poptávkovou $C_D(R)$ (vzhledem k pracovišti):

$$C_S(R) = \frac{\sum_{i \in R, j \in R} f_{ij}}{\sum_{i \in R} o_i}$$

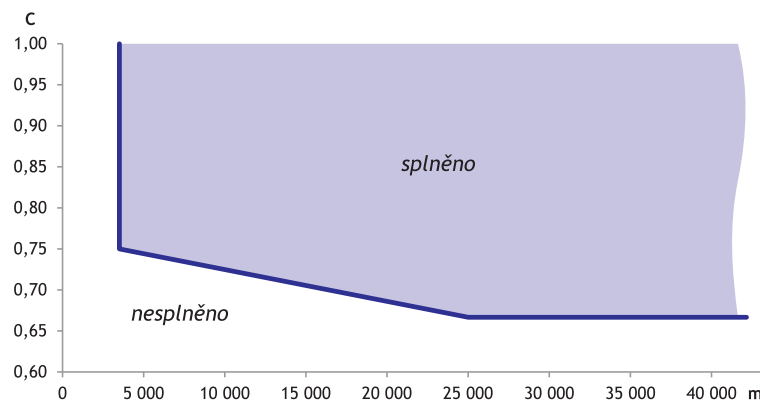
$$C_D(R) = \frac{\sum_{i \in R, j \in R} f_{ij}}{\sum_{i \in R} i_i}$$

Vlastní kritérium má potom formu

$$C_0 = C_1 + \frac{C_2 - C_1}{m_2 - m_1} (\min\{m, m_2\} - m_1),$$

kde m je masa regionu a $m_1 = 3500$, $m_2 = 25000$, $C_1 = \frac{3}{4}$, $C_2 = \frac{2}{3}$ empiricky stanovené parametry, určující požadavky na velikost a uzavřenost regionu, přičemž tyto dva požadavky se mohou vzájemně kompenzovat (viz obr. 7). Hranice je stanovena jako $\min\{C, C_1\} > C_0$, kde $C = \min\{C_S, C_D\}$ je uzavřenost regionu.

Obrázek 7: Coombesovo ověřovací kritérium a jeho hranice. Šikmý segment hranice naznačuje vzájemnou kompenzaci (trade-off) mezi uzavřeností C a masou m regionu.



Na začátku běhu je každá zóna považována za samostatný region. V každém kroku je rozpuštěn region s nejnižší hodnotou ověřovacího kritéria a jeho zóny jsou dle agregačního kritéria přiřazeny ostatním regionům, a to tak dlouho, dokud existuje region, který ověřovací kritérium nesplňuje.

Metoda neobsahuje podmínku sousedství s odůvodněním, že by běh metody výrazně deformovala; místo toho jsou exklávy eliminovány manuálně *ex-post* na základě subjektivního rozhodnutí.

Starší verze metody (Coombes a kol. 1982) obsahovala slučovací kritérium na bázi

agregačního kritéria. Dvě jádra byla na začátku běhu algoritmu sloučena v jeden vícejaderný region, pokud byl jejich vzájemný vztah dostatečně silný ($s_{ij2} \geq \frac{1}{100} \wedge f_{ij} \geq \frac{1}{10} o_i \wedge f_{ji} \geq \frac{1}{100} o_j$). Také bylo aplikováno přerazovací kritérium pro lokální optimalizaci s_{ij2} na hraniční zóny regionů.

Ve starší verzi je také kromě výše uvedeného kritéria s_{ij2} diskutována jeho odlišná forma

$$s_{ij3} = \frac{f_{ij}}{o_i} + \frac{f_{ij}}{i_j} + \frac{f_{ji}}{o_j} + \frac{f_{ji}}{i_i}.$$

Autoři – a dle nich i Bezák (2000) – argumentovali tím, že Smartova míra s_{ij2} ve slaběji propojených systémech (zejm. v řidčeji osídlených oblastech) neadekvátně zvýhodňuje menší centra před většími díky nižším hodnotám marginálních součtů. Aplikace novější verze metody CURDS, kde je použita míra s_{ij2} , na český sídelní systém (Klapka a kol. 2013) ukazuje, že tato kritika je oprávněná.

2.6.4 Bezákova regionalizace Slovenska

Bezák (2000) použil upravenou starší verzi metody CURDS. Zde je tento pokus zmíněn kvůli své propracovanosti a relativní podobnosti slovenského sídelního i statistického systému s českým.

Selekční kritérium obsahuje dva lokalizační kvocienty – jeden vztahuje počet pracujících v zóně k ekonomicky aktivnímu obyvatelstvu zóny (a zohledňuje tak pracovní funkci), druhý počet pracujících ve službách a prostý počet obyvatel zóny (čímž zohledňuje obslužnou funkci). Jejich minimální hodnotu pro jádro stanovuje na 0,7.

Ověřovací kritérium používá jinou míru uzavřenosti, která je dle autora upravená tak, aby kompenzovala chybějící data o vnitrozónálních a málo významných tocích:

$$C(R) = \frac{\sum_{i \in R, j \in R} f_{ij}}{\sum_{i \in R, j \notin R} f_{ij} + \sum_{i \notin R, j \in R} f_{ij}}$$

Z této míry a požadavku minimální velikosti je konstruováno samo kritérium:

$$Z(R) = \min \{1, C(R)\} \cdot \min \left\{1, 1,0526 \cdot \frac{m(R)}{m_0}\right\},$$

kde $m(R)$ je váha regionu a m_0 je minimální požadovaná váha; toto kritérium neumožňuje vzájemnou kompenzaci velikosti a uzavřenosti.

Slučovací kritérium je podstatně restriktivnější než u metody CURDS, avšak mimo kooperujících center se zde objevuje snaha sloučit do jednoho také regiony center, které si vzájemně konkurují, a to v případě, že nevyhovují ověřovacímu kritériu a platí, že:

$$\frac{\sum_{i \in R_1, j \in R_2} f_{ij}}{\sum_{i \in R_1, j \notin R_1} f_{ij}} \geq 20\% \wedge \frac{\sum_{i \in R_2, j \in R_1} f_{ij}}{\sum_{i \in R_2, j \notin R_2} f_{ij}} \geq 10\%$$

2.6.5 Metoda funkční vzdálenosti

Koncept funkční vzdálenosti pro vymezení funkčních regionů poprvé použili Brown a Holmes (1971); v novější literatuře se jím zabývá Cörvers a kol. (2009). Vychází ze zásady zohlednění nepřímých toků, pro což využívá *markovské řetězce*. Pomocí tohoto matematic-

kého aparátu, kdy imaginární cestovatel prochází zónami jako jednotlivými stavy s pravděpodobností přechodu na danou zónu rovnou síle interakce, která do ní směřuje z aktuální zóny, transformuje interakční matici na matici *středních dob prvního průchodu* (*mean first passage time, MFPT*), které představují funkční vzdálenosti mezi zónami.

Vlastní seskupení je potom prováděno hierarchickou procedurou, která slučuje zóny s nejnižší funkční vzdáleností za podmínky sousedství (ta je zde nezbytná, neboť MFPT dostatečně nezohledňuje relativní význam zóny). Procedura se snaží dosáhnout předem *a priori* stanoveného preferovaného počtu regionů, jehož stanovení je zákonitě problematické (Cörvers a kol. je stanovují rovný počet administrativních regionů příslušné úrovně).

2.6.6 Evoluční metody

V poslední době se objevil též pokus o vymezování funkčních regionů stochastickým postupem na bázi evolučních algoritmů (Martínez-Bernabeu a kol. 2012). Evoluční algoritmus požaduje dané vymezení regionů za jedince, jehož kvalita je dána ověřovacími kritérii; tyto jedince mezi sebou kříží (operátor křížení), případně vytváří nové mutací (skupinové mutační operátory).

Na začátku je vytvořena populace výchozích jedinců pomocí algoritmu SHA, který provádí postup podobný Coombesově regionalizaci, ovšem místo exaktních kritérií pracuje s pravděpodobnostmi, takže zóna má větší pravděpodobnost přiřazení k regionu, s nímž silněji interaguje.

Vlastní běh algoritmu potom spočívá v cyklech. V každém cyklu je vytvořeno několik nových jedinců pomocí mutačních a křížicích operátorů, ti jsou seřazeni dle ověřovacího kritéria a do dalšího cyklu postupuje pouze určený počet nejlepších. Cyklus se opakuje tak dlouho, dokud není stabilně obsazena pozice nejlepšího jedince.

Křížicí operátor vybírá náhodně dva jedince (s pravděpodobnostmi rostoucí s jejich kvalitou) a náhodně zkopíruje vymezení regionů z jednoho nebo druhého. Mutačních operátorů je devět, pracujících na podobném principu. Pokrývají téměř veškeré myslitelné operace s regiony (sloučení dvou regionů, rozdělení regionu na dva, vytvoření nového regionu ze zóny jiného, přearování zón mezi regiony).

Po celou dobu běhu je zachována podmínka sousedství; autoři však poznamenávají, že její omezující efekt je minimalizován stochastickým charakterem algoritmu a naopak umožňuje výrazně snížit počet procházených variant, čímž se běh výrazně zrychluje oproti předcházející variantě (Flórez-Revuelta a kol. 2006).

Autoři tvrdí, že jejich algoritmus vymezuje v porovnání s metodou CURDS výrazně kvalitnější funkční regiony (vzhledem k definovanému ověřovacímu kritériu, jež mají obě metody velmi podobné). Od evolučních algoritmů a jejich dalšího vylepšování lze skutečně očekávat velký přínos zejména pro realistické pojetí vymezování funkčních regionů.

Nevýhodou evolučních algoritmů je však právě jejich nedeterministický charakter, který nezaručuje při stejných vstupních datech totožný výsledek; rovněž absence přímo definovaného jasného postupu může být na škodu při obhajování objektivit regionalizace pro pragmatické pojetí.

2.6.7 Shrnutí

Z výše uvedeného je zřejmé, že k vymezení funkčních regionů existují velmi rozdílné přístupy co do postupů, ale v některých kritériích jsou si podobné. Většina přístupů se snaží do vymezení zahrnout i prostorovou dimenzi (nejčastěji pomocí sousedství), a proto by mohly profitovat z implementace v GIS. Žádný z přístupů nalezených v české nebo světové literatuře však GIS nevyužívá (FlowMap, v němž je implementován Intramax, je pouze úzce specializovaným programem). Tento deficit se tato práce snaží napravit. Pokouší se přitom o syntézu výše představených metod, přičemž z nich vybírá z nich ty prvky, které lze v GIS vhodně použít.

3 Návrh metody vymezení

V této části je navržen vlastní regionalizační algoritmus primárně vycházející z metody, kterou použili Hampl a Marada (2014), ovšem zahrnující v sobě vlivy ostatních prací (Bezák 2000, Coombes a kol. 1986, van der Laan a Schalke 2001, Watts 2013) a přidávající některé prvky specifické pro GIS implementaci. Hamplova metoda je jako základní stavební kámen zvolena především díky své teoretické propracovanosti a dlouhé tradici kvalitních výsledků v českém sídelním systému. Některé její koncepty jsou v rámci světové literatury unikátní a neprávem opomíjené.

Algoritmus začíná se stavem, v němž je každá zóna jádrem svého vlastního regionu. Regiony jsou pak seřazeny podle ověřovacího kritéria a ty, které jej nesplňují, jsou postupně eliminovány (jejich zóny jsou přiřazeny ostatním regionům nebo je celý region sloučen s jiným za vzniku regionu vícejaderného).

Na konci je potom provedena optimalizace výsledku přerazováním okrajových zón. Přesná podoba algoritmu je znázorněna v diagramu na obrázku 8; níže jsou potom popsána použitá kritéria.

3.1 Selekční kritérium

Selekční kritérium není na rozdíl od Hamplovy metody aplikováno v souladu s myšlenkou, že regionální jádra by měla být vybírána právě na základě schopnosti udržet si dostatečné zázemí (viz s. 14) – jde tedy o induktivní metodu. (Pro zrychlení běhu lze selekční kritérium aplikovat ve značně uvolněné a jednoduché formě, která vyřadí pouze jasně necentrální místa – pro Česko se jeví jako vhodná spodní populační velikost 1 000 obyvatel.)

3.2 Agregáčn  krit rium

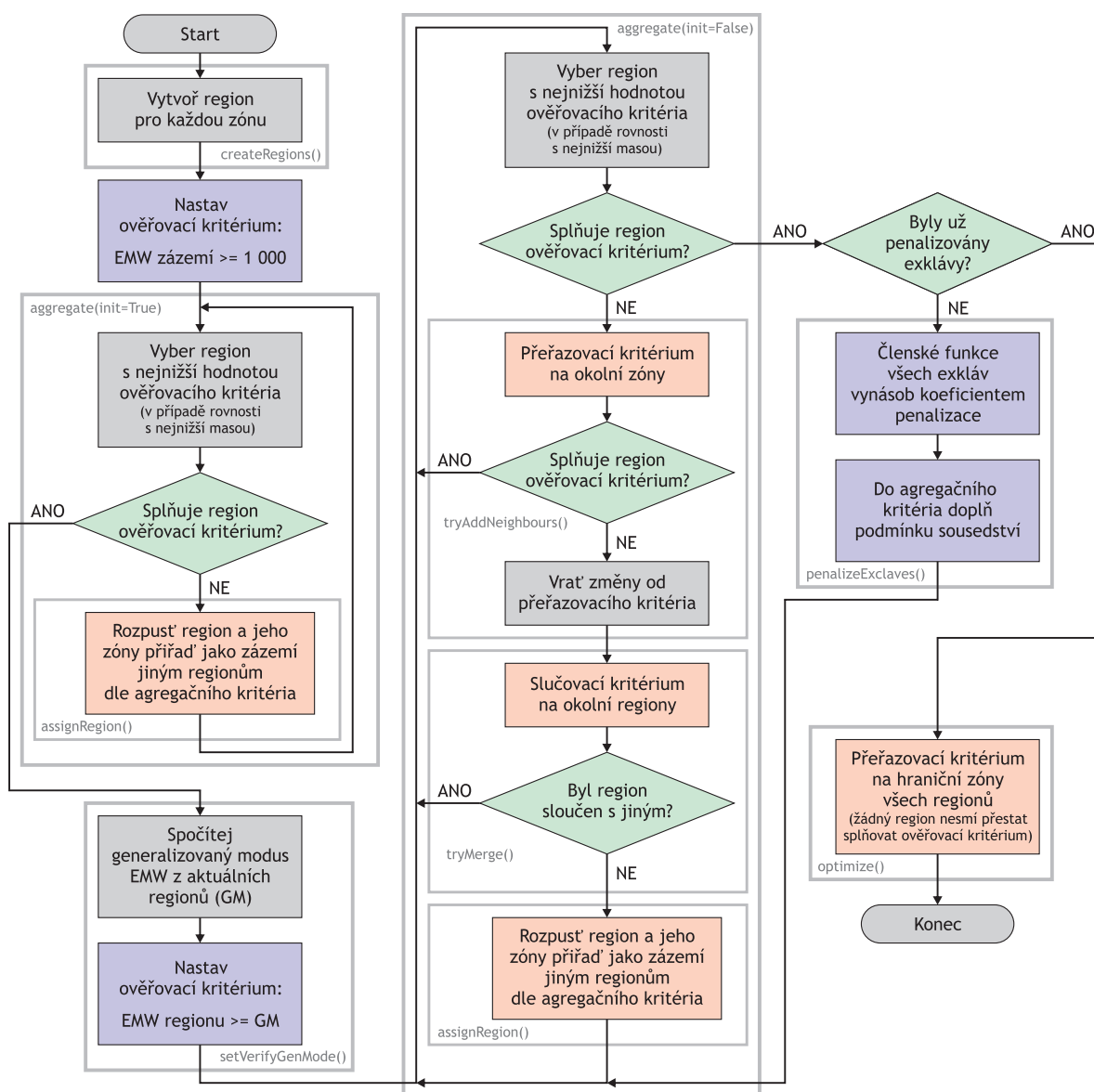
Agregační kritérium a_{ij} je založeno na Coombesově míře s_{ij3} (viz s. 22) upravené na vztah jednosměrné závislosti do podoby:

$$a_{ij} = \frac{f_{ij}}{o_i} + \frac{f_{ij}}{i_j},$$

kde f_{ij} je síla interakce ze zóny i do zóny j , o_i součet síly všech výchozích toků ze zóny i a i_j součet síly všech příchozích toků do zóny j . Kritérium vychází z míry zohledňující význam toku pro výchozí i cílovou zónu, z níž je ovšem v souladu s Hamplovou metodou vyloučena síla zpětné interakce. Vícečlenná jádra regionů jsou přitom považována za jednu zónu (jako cíl i jako počátek – při rozpouštění regionu jsou agregována společně). Při rozpouštění regionu je nejprve přiřazeno jádro a poté zázemí.

Použitá forma agregáčn ho krit ria preferuje interakce s menšími centry (díky druh mu  lenu) tak, aby se v počátku mohly vytvořit i menší regiony; selekci životaschopných regionů

Obrázek 8: Vývojový diagram navrhovaného algoritmu. Jednotlivá kritéria jsou vysvětlena níže.



ponechává na ověřovacím kritériu a finální podobu hranic regionů (přiřazení okrajových zón) na přeřazovacím kritériu. Zároveň je ovšem použitím míry s_{ij3} místo s_{ij2} zabráněno neadekvátní preferenci menších center.

Pokud takto určený nejsilnější tok ($\max_j a_{ij}$) nesměruje do jádra, ale do zóny zázemí k , je aplikováno *pravidlo nepřímé vazby* a interakční vektor \vec{f}_i je transformován:

$$\vec{f}_i = \vec{f}_i + \frac{f_{ik}}{o_i} \vec{f}_k$$

Je pak velmi pravděpodobné, že zóna i bude přiřazena ke stejnému regionu jako zóna k , ovšem pokud má dostatečně silný sekundární tok, který se neshoduje s primárním tokem zóny k , mohou obě zóny skončit v různých regionech.

Pravidlo nepřímé vazby je motivováno snahou, aby co možná nejvíce zón agregovaného regionu zůstalo pohromadě a byla tak zachována zásada skladebnosti nižších úrovní dle M. Hampla (viz s. 18). Region však na rozdíl od Hamplova přístupu není agregován jako celek – je tak umožněno zónám, které připadly původnímu agregovanému regionu jen těsně, aby byl místo ještě slabší vazby na nový region využit jejich sekundární tok.

V závěrečné fázi algoritmu je agregační kritérium omezeno *podmínkou sousedství* – zóna se nesmí stát exklávou.

3.2.1 Zahnutí vzdáleností

Interakce mohou být volitelně před vstupem do algoritmu transformovány gaussovskou funkcí vzdálenosti d_{ij} (Farmer a kol. 2011, vlastní úprava)

$$f'_{ij} = f_{ij} \cdot \exp\left(\frac{-d_{ij}^2}{k_d b}\right),$$

kde b je určeno na základě vstupního souboru metodou OLS (viz níže) a k_d je parametr, jehož hodnota určuje přísnost eliminace interakcí s nadprůměrně vysokou vzdáleností – u nich se předpokládá, že integrují vyšší řádovostní úroveň regionů, než je právě vymezována, a proto by do výsledku pouze zanášely chybu.

Hodnota parametru b úbytku síly interakce se vzdáleností je určena na základě analýzy vstupního souboru interakcí. S rostoucí vzdáleností se totiž množství i síla interakcí zmenšují. Vstupní interakce jsou proto rozděleny na n kvantilů dle vzdálenosti d_{ij} a v rámci těchto kvantilů je síla všech interakcí sečtena (aby byl zahrnut jak počet interakcí na danou vzdálenost, tak jejich síla) do součtu F_k . Pro každý kvantil Q_k je také spočítána průměrná interakční vzdálenost D_k :

$$\begin{aligned} Q_k &= \langle d_k, d_{k+1} \rangle \\ F_k &= \sum \{f_{ij} | d_{ij} \in Q_k\} \\ D_k &= \frac{\sum \{d_{ij} | d_{ij} \in Q_k\}}{|\{d_{ij} | d_{ij} \in Q_k\}|} \end{aligned}$$

Parametr b je potom spočítán optimalizací metodou nejmenších čtverců – minimalizací

výrazu

$$\sum_{k=1}^n \left(F_k - g \cdot \exp \left(\frac{-D_k^2}{b} \right) \right)^2,$$

kde g je převodový faktor (převádí exponenciální člen na jednotky interakce).

3.3 Ověřovací kritérium

Ověřovací kritérium je založeno na *členských funkcích*, jež použil Feng (2009) a zdokonalil Watts (2009) pro určování příslušnosti zóny k regionu.

3.3.1 Členské funkce

Členská funkce určuje míru integrace zóny do regionu a nabývá hodnot $d_i \in \langle 0; 1 \rangle$, kde 0 odpovídá neexistenci vazby zóny na region, 1 pak vazbě absolutní. Feng (2009) počítá členskou funkci (zobecněně) jako podíl toků zóny integrujících daný region ku všem tokům zóny.

Zde použitá členská funkce je založená na Hamplově pojetí regionálních vazeb (viz obr. 5 na s. 19). Ta považuje za integrující pouze toky mezi jádrem a zázemím regionu, nikoliv toky v rámci zázemí – ty jsou proto z kalkulace úplně vyňaty. Přímou se tak vychází z Hamplových měř integrity.

Pro zónu v zázemí regionu $i \in h(R)$ je funkce členství v regionu R vyjádřena jako:

$$d_i(R) = k_e \frac{I_{iR}}{I_{iR} + D_{iR}},$$

kde $I_{iR} = \sum_{j \in c(R)} f_{ij}$ je tok ze zóny do jádra regionu (považovaný za integrující) a $D_{iR} = \sum_{j \notin R, j \neq i} f_{ij}$ součet toků ze zóny mimo daný region (považovaných za dezintegrující).

k_e je potom *koefficient sousedství*; pro přiřazení zóny k regionu, s nímž zóna sousedí⁹, má hodnotu 1, pro přiřazení k nesousednímu regionu má hodnotu *koefficientu penalizace exkláv* $k_{e0} \in \langle 0; 1 \rangle$, která určuje snížení hodnoty takového přiřazení. Hodnota k_{e0} je na začátku algoritmu 1 (žádný rozdíl), v závěrečné fázi algoritmu je však nastavena na 0. To odpovídá snaze nejdříve vymežit regiony na základě interakcí (bez ohledu na souvislost) a teprve potom je přerazovacím kritériem upravit tak, aby splňovaly princip souvislosti – v souladu s Hamplovým přístupem, ovšem převedeným z manuální do algoritmické formy.

Pro zónu v jádru regionu $j \in c(R)$ je členská funkce stanovena odlišně, a to ze dvou důvodů:

- Míra integrace jádra do regionu závisí nejen na jeho odchozích, ale i na jeho příchozích tocích. Proto se pro potřeby kalkulace členské funkce jádrové zóny oba směry interakcí sčítají.
- Za integrující toky se považují všechny interakce zóny jádra v rámci regionu, jak do zázemí, tak do ostatních zón jádra.

⁹Sousední region je takový, že ze zóny existuje cesta k alespoň jednomu jádru daného regionu pouze skrz jeho zóny. Zóna tak s regionem nemusí sousedit ani v případě, kdy sousedí s jednou z jeho zón; tato zóna může totiž být exklávou.

Členy I_{jR} , D_{jR} pro jádrovou zónu tedy nesou podobu:

$$I_{jR} = \sum_{k \in R, k \neq j} f_{jk} + f_{kj}$$

$$D_{jR} = \sum_{k \notin R} f_{jk} + f_{kj}$$

3.3.2 Členská váha

Pomocí členských funkcí je možné definovat *fuzzy členskou váhu regionu (FMW)* jako

$$W_f(R) = \sum_{i \in Z} m_i d_i(R),$$

která udává celkový význam centra pro své okolí. Výhodou je, že fuzzy váha v sobě automaticky obsahuje velikost i uzavřenost, takže ověřovací kritérium mohou splnit jak menší regiony se silnými vazbami jádro-zázemí, tak větší regiony s nižší integritou. Další (byť diskutabilní) výhodou je možnost vytvořit region, který má sice malé zázemí orientované na něj primárním tokem, ovšem vykazuje vysokou sekundární spádovost.

Pro splnění principu jednoznačnosti lze využít *exkluzivní členskou váhu regionu (EMW)*:

$$W(R) = \sum_{i \in R} m_i d_i(R),$$

kde nejsou uvažovány vazby těch zón, které do regionu nepatří.

3.3.3 Stanovení hranic

Pro úplné stanovení ověřovacího kritéria je nutno kromě ukazatele také vhodně zvolit jeho minimální hodnotu. Ukazuje se, že právě velikostní hranice jsou nejčastěji upravovaným parametrem při přejímání regionalizační procedury, a to kvůli adaptaci na odlišný sídelní systém; to ovšem zároveň znemožňuje věrohodné srovnání mezi studiemi. Nejvýhodnějším (byť zatím zřídka využívaným) přístupem je použít hodnotu získanou na základě analýzy vlastního sídelního systému – tak je postupováno i zde. Hranice minimální velikosti je stanovena pomocí Hamplovy techniky generalizovaného modu (viz obr. 6 na s. 19), a to výpočtem z těch regionů, které mají nezanedbatelné zázemí (postup výpočtu je uveden v části 4.2.4 na straně 39). Proto jsou v algoritmu použita dvě ověřovací kritéria:

1. Hodnota je rovna EMW zázemí regionu. Minimální hranici určuje parametr W_0 , jehož hodnota by měla být dostatečně malá, aby nebyly zbytečně likvidovány i regiony, které by jinak splňovaly druhé ověřovací kritérium; zároveň však dostatečně velká pro formování alespoň minimální polarity jádro-zázemí. Jeho určení je proto ponecháno sociogeografické analýze zkoumaného sídelního systému.
2. Hodnota je rovna EMW celého regionu. Požadována je hodnota rovná generalizovanému modu EMW regionů splňujících předchozí kritérium (alternativně lze hodnotu stanovit pevně – W_1).

Dvoustupňové agregační kritérium vyžaduje také dvoustupňovou agregaci, jak je vidět i z vývojového diagramu na obr. 8. Pokud platí první ověřovací kritérium, není na eliminované

regiony před agregací zkoušeno přeražení ani sloučení, neboť se v této fázi vyskytují i regiony s nulovou EMW (tvořené jednou zónou a tudíž nemající vnitřní toky); takové regiony by získávaly při přerazování a slučování výhodu neodpovídající jejich skutečnému významu, a proto jsou rovnou agregovány.

3.4 Přerazovací kritérium

Přerazovací kritérium r_i je postaveno na bázi EMW regionu. Pro přeražení zóny i z regionu A do regionu B má hodnotu

$$r_i(A \rightarrow B) = \frac{\Delta W(B)_i - \Delta W(A)_i}{m_i} > r_0,$$

kde $\Delta W(A)_i$ je změna EMW regionu A vyvolaná přeražením a m_i masa zóny i . Hodnota r_0 (*mez přiražení*) je malé kladné číslo, zabraňující provedení změn s příliš malým dopadem vzhledem k velikosti zóny. Pokud je r_i větší než tato mez, přeražení je provedeno.

Přerazovací kritérium je po celou dobu běhu algoritmu omezeno podmínkou sousedství v tom smyslu, že přerazovaná zóna se nesmí stát exklávou; zde jde především o snahu omezit počet prohledávaných variant. V závěrečné fázi (po agregaci všech regionů nesplňujících ověřovací kritérium) zároveň nejsou povolena ta přeražení, která by jeden z regionů posunula pod hranici splnění ověřovacího kritéria.

V původní metodě M. Hampla se přerazovací kritérium vyskytuje pouze pro přeražení exkláv; zde je rozšířeno s cílem optimalizovat výslednou podobu regionů tak, aby se pro každou zónu maximalizovala její členská funkce a tím bylo učiněno zadost principu autonomie.

Výpočet změny EMW vyvolané přeražením je pro přeražení zóny i do regionu R určen vzorcem

$$\Delta W(R)_i = m_i d_i(R) + \sum_{j \in R} m_j \Delta d_j(R, i),$$

kde $\Delta d_j(R, i)$ je změna členské funkce zóny j v regionu R vyvolaná přeražením zóny i :

$$\Delta d_j(R, i) = \frac{I_{jR}}{I_{jR} + O_{jRi}} - e \cdot \frac{I_{jR}}{I_{jR} + O_{jRi} + f_{ji}},$$

kde $O_{jRi} = \sum_{k \notin R, k \neq i} f_{jk}$ je tok ze zóny j do všech zón vně regionu R kromě zóny i a $e = k_{e0}$ v případě, že přeražení zóny i změní souvislost¹⁰ zóny j s regionem R , jinak $e = 1$.

3.5 Slučovací kritérium

Slučovací kritérium je povoleno pro sloučení dvou regionů A, B se silnou vzájemnou vazbou a přitom nízkou masou. Je použit ukazatel relativního překryvu $RO(A, B)$ (Watts 2013, upraveno):

$$RO(A, B) = \frac{\sum_{i \in A} m_i d_i(B) + \sum_{j \in B} m_j d_j(A)}{W(A) + W(B)} > RO_0,$$

kde RO_0 je hraniční hodnota určující pravděpodobnost tvorby vícejaderných regionů.

¹⁰Tedy při přeražení zóny i se zóna j stane exklávou, pokud jí není, a naopak exklávou být přestane, pokud jí právě je.

Cílem aplikace slučovacího kritéria je zohlednit situace v území, kdy některá blízká sídla vykazují vysoký stupeň územní dělby práce a vzájemné kooperace, čímž formují jádro jediného funkčního regionu. M. Hampl takové situace určuje na základě odborné úvahy nad konkrétní situací.

3.6 Sousedství

Sousedství je v této metodě založeno na dopravních sítích. Dvě zóny spolu sousedí, pokud mezi jejich sídly existuje dopravní spojení neprocházející žádným sídlem, které nepatří do jedné z obou zón. Alternativně lze využít jiný koncept sousedství z části 2.5.1.

Podmínka sousedství se promítá do agregačního a přeřazovacího kritéria a skrz koeficient penalizace exkláv i do kritéria ověřovacího; rozhodující úlohu ale hraje až v závěrečné fázi algoritmu. Smyslem tohoto odkladu je nebránit vzniku organických jednotek dle hodnot interakcí a teprve potom jejich výsledky smysluplně (nikoli *ad hoc*) korigovat, aby jednotky vyhovovaly i principu souvislosti.

3.7 Parametry metody

Ač byla metoda navrhována tak, aby potřebovala pouze nezbytné minimum parametrů určených zvnějšku, přesto to je u několika hodnot nutné – nejlépe na základě analýzy zkoumaného sídelního systému a určených priorit jednotlivých principů vymezení s ohledem na pojetí regionalizace, která je zpracovávána (viz část 2.3)¹¹. Těmito parametry jsou:

- Hranice EMW pro úvodní ověřovací kritérium W_0 (viz část 3.3.3, s. 29).
- Koeficient penalizace exkláv k_{e0} , určující tvrdost aplikace principu souvislosti od 0 (povoleny pouze souvislé regiony) po 1 (na souvislosti regionů nezáleží).
- Minimální přeřazovací zisk r_0 (v jednotkách masy). Jeho použití je nutné, aby nedocházelo k příliš velkému množství přeřazení, které běh programu příliš zpomalují a přitom nemají žádný významný efekt.
- Minimální slučovací překryv $RO_0 \in (0; +\infty)$, který určuje, při jak velkém překryvu regionů, z nichž alespoň jeden nesplňuje ověřovací kritérium, má dojít k jejich sloučení a tvorbě regionu vícejaderného; hodnota blízká nule pravděpodobně vytvoří velké množství vícejaderných regionů, hodnoty přesahující 1 většinu sloučení zabrání. Stanovení algoritmu pro výpočet tohoto parametru ze vstupních dat se ukázalo jako úkon přesahující rozsah této práce.
- Koeficient eliminace interakcí s nadprůměrnou vzdáleností $k_d \in (0; +\infty)$ (pokud je aplikována vzdálenostní korekce). Hodnota blízká nule znamená rapidní pokles síly se vzdáleností, velmi vysoké hodnoty potom pokles téměř nezatelný.
- Volitelně lze určit i přesnou hranici EMW regionu W_1 pro hlavní ověřovací kritérium – např. v případě, že pro daný regionální systém nedává technika generalizovaného modu smysluplný výsledek.

¹¹Pro co nejsnazší adaptaci metody na různé sídelní systémy je většina parametrů relativních.

3.8 Vizualizace výsledků

Členské funkce zón umožňují snadno určit míru, s jakou zóna patří různým regionům – a tuto míru lze snadno vizualizovat pomocí barvy. Každému regionu přiřadíme určitou sytou barvu $C(R)$ a tyto odstíny pro jednotlivé zóny i smísíme podle vzorce

$$C_i = \sum_{(I_{iR} \neq 0)} C(R) \cdot d_i(R),$$

kde $d_i(R)$ je členská funkce. Pro jednoduchost reprezentace a manipulace je použit barevný model RGB – zóny se silnou vazbou k jádru svého regionu budou tak nést jeho sytou barvu, zóny se silnou vyjížděkou mimo jádra se budou vyznačovat nízkou sytostí barvy (blízkou šedé) a oscilační území s významnými vazbami na jádra více regionů budou mít barvu velmi světlou. Pro sofistikovanější kartografické znázornění by samozřejmě bylo vhodnější použít jiný model (např. HSV; více viz např. Brown a Feringa 2003).

4 Aplikace

Tato část se věnuje implementaci výše představených metod do podoby geoprocessingových nástrojů umožňujících provádět příslušné analýzy přímo v prostředí GIS. Pro implementaci bylo vybráno rozhraní ArcPy programu ArcGIS, umožňující snadnou tvorbu podobných nástrojů na bázi skriptů v jazyku Python.

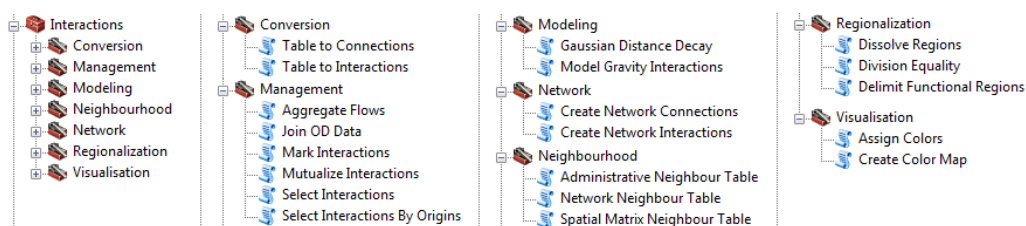
ArcGIS, ač komerční program, byl zvolen především díky relativně jednoduchému rozhraní, a to jak uživatelskému (snadná tvorba *toolboxů* obsahujících jednoduché grafické uživatelské rozhraní) i programátorskému (Python umožňuje vysokou flexibilitu); dále pak díky široké funkcionalitě i rozšíření mezi uživateli. Daní za jeho použití je poněkud nižší výkon skriptů oproti např. QGISu a vysoká cena programu. Většina tříd nástroje je proto napsána tak, aby byla nezávislá na použitém GISu; případná konverze do jiného prostředí tak zahrnuje pouze přepsání několika tříd využívajících přímo funkci modulu `arcpy` – `ZoneLoader`, `InteractionLoader`, `NeighbourLoader`, `GeometricalZone` a moduly `common` a `networking`.

V rámci této práce byl vytvořen toolbox **Interactions**, který zajišťuje uživatelské rozhraní k přiloženým skriptům v jazyce Python. Toto uživatelské rozhraní umožňuje pohodlné zadávání vstupních dat a parametrů přímo z prostředí ArcGIS včetně jejich jednoduché validace. V této kapitole je nejprve stručně představena funkcionalita tohoto toolboxu, který je přiložen na CD, způsob implementace jeho vybraných částí a nakonec předpokládaný postup jeho použití (část 4.4). Následně je představena ukázka aplikace – vymezení funkčních regionů dle dat ze Sčítání lidu, domů a bytů 2011.

4.1 Struktura a prvky toolboxu

Vytvořený toolbox je zobrazen v grafickém rozhraní ArcGIS na obrázku 9 a přiložen na CD. Skládá ze sedmi toolsetů, jež jsou popsány níže. O samotné vymezení funkčních regionů se stará toolset **Regionalization** na základě metody popsané v kapitole 3, zbylé toolsety obsahují nástroje pro přípravu potřebných vstupních dat a vizualizaci výsledků:

Obrázek 9: Nástroje ve vytvořeném toolboxu tak, jak se zobrazují v grafickém rozhraní ArcGIS.



- **Conversion** umožňuje automaticky georeferencovat interakce – jeho nástroje z tabulky interakcí vytvoří liniovou vrstvu na základě bodové vrstvy interagujících zón.

Table to Interactions tvoří linie vzdušnou čarou, **Table to Connections** cesty (*routes*) po zadané síti (*network datasetu*). Podrobněji jsou popsány v části 4.3.2.

- **Management** zahrnuje pomocné nástroje pro manipulaci s interakcemi:
 - **Aggregate Flows** agreguje interakce mezi zónami do interakcí mezi vymezenými regiony,
 - **Join OD Data** přiřazuje interakcím atributy jejich výchozí a cílové zóny,
 - **Mark Interactions** určuje interakcím jejich relativní význam pro výchozí zónu,
 - **Mutualize Interactions** tvoří z nesymetrických interakcí symetrické,
 - **Select Interactions** vybírá interakce dle síly a vzdálenosti a
 - **Select Interactions By Origins** omezuje interakce pouze na určitou podmnožinu zón (vybírá takové, jež probíhají mezi zónami z vybrané vrstvy).
- **Modeling** zahrnuje nástroje modelující sílu interakce dle gaussovského vzdálenostního modelu (**Gaussian Distance Decay**, více viz část 4.3.2) nebo gravitačního modelu (**Model Gravity Interactions**). Tyto nástroje potřebují pro svůj chod v Pythonu nainstalovaný matematický balík **scipy** (SciPy 2014; volně stažitelný).
- **Network** obsahuje nástroje pro hromadné generování liniových interakcí ze zadané sítě (*network datasetu*) s údaji o vzdálenosti, resp. jiných nákladových atributech. **Create Network Interactions** využívá analýzu **OD Matrix**, a proto generuje pouze přímé linie, byť s přesnými hodnotami nákladových atributů; **Create Network Connections** využívá **Closest Facility**, je proto pomalejší, ale zato generuje přesné linie cest.
- **Neighbourhood** generuje tabulky sousedství pro vstup do regionalizace; více o jednotlivých nástrojích je uvedeno v části 4.3.1.
- **Regionalization** obsahuje jádro toolboxu – nástroje vymezující funkční regiony:
 - **Delimit Functional Regions** vymezuje funkční regiony a ukládá informace do atributové tabulky zón. Je podrobně popsán v části 4.2.
 - **Dissolve Regions** podle této informace generuje polygony regionů s agregátními informacemi o jádře i celém regionu.
 - **Division Equality** umožňuje spočítat shodnost dvou různých vymezení uložených v atributové tabulce zón (počítá, jaký podíl z masy všech zón je oběma vymezeními zařazen do shodného regionu).
- **Visualisation** umožňuje vizualizaci příslušností zón k regionům dle postupu popsaného v části 3.8 na s. 32. Technické detaily jsou uvedeny v části 4.3.3.

Pro více informací o technických detailech použití nástrojů viz jejich dokumentaci, jež se zobrazí po otevření nástroje v ArcCatalogu nebo ArcMapu.

Každý nástroj je tvořen skriptem, uloženým jako `.py` soubor. Kromě nich patří k toolboxu ještě tři moduly – **common**, obsahující jednoduché funkce pro práci s rozhraním ArcPy, **objects**, obsahující třídy pro objekty využívané regionalizačními nástroji a **networking**, pokrývající konverzi interakcí z tabulek na geometrické vrstvy, ať už přímo, nebo pomocí síťové analýzy. Právě tyto tři moduly obsahují většinu funkcionality; soubory skriptů po většinou pouze importují moduly a volají funkce jejich objektů.

Většina nástrojů má povoleno přepisování (`arcpy.env.overwriteOutput = 1`).

4.2 Implementace vymezení metody

Vlastní metoda vymezení funkčních regionů popsaná spolu s vývojovým diagramem v kapitole 3 je implementována v nástroji **Delimit Functional Regions (DFR)** z toolsetu **Regionalization**. Zde je představena struktura tříd, které nástroj využívá, požadovaná vstupní data a parametry a některé prvky jeho implementace.

Protože program pracuje primárně se sociogeografickými daty, jež jsou ze své povahy diskrétní, je použit *objektový (object-based)* přístup (Worboys a Duckham 2004), jež vnímá jednotlivé jevy jako objekty a jejich vlastnosti, oproti *field-based* přístupu, jež považuje jevy za vlastnosti prostoru. Objektový přístup potom vede k užití vektorových dat.

4.2.1 Struktura tříd

Struktura tříd, jež nástroj využívá, je popsána na obrázku v příloze 2. Jádru tvoří třídy představující zóny (**Zone**, které jsou podle selekčního kritéria buď instancí potomka **CoreZone**, nebo **HinterlandZone**), jejich interakce (**Interactions** – slovník cíl : síla toku obsahující všechny interakce daného objektu směřující z něj nebo do něj), regiony (**Region**) a přiřazení zón k regionům (**Assignment**, obsahující hodnotu členské funkce a umožňují vztah efektivně spravovat).

Datový model byl v průběhu implementace několikrát přepracováván ve snaze najít ten nejobecnější (ve smyslu použitelnosti na mnoho úloh) a přitom nejefektivnější. Třída **Assignment** byla zařazena, aby byla umožněna i implementace fuzzy metod, kdy je jedna zóna členem více regionů zároveň. Interakce by mohly rovněž být reprezentovány po jedné jako objekt s počátkem, cílem a silou, avšak tento přístup byl zavržen pro svou vysokou výpočetní náročnost; zvolená implementace pomocí potomka **collections.defaultdict** je velmi programátorsky komfortní a umožňuje přímo v sobě provádět množství výpočtů (souhrny interakcí do daného regionu ap.).

Třídy **Regionaliser**, **Bindings** a **NeighbourChange** se snaží oddělit aplikaci vymezení kritérií konkrétní metody od datových tříd. To se zcela nepodařilo; datové třídy obsahují část funkcionality specifické pro vlastní vymezení metody (byť to nevylučuje jejich použití v implementaci jiné vymezení metody).

4.2.2 Vstupní data a parametry

Vstupní data nástroj přijímá ve třech tabulkách:

- Tabulku zón (**Zones**), obsahující veškeré informace o zónách ovlivňující provedení regionalizace. Nástroj z tabulky čerpá především identifikátory zón (**Zone ID Field** – numerické nebo textové) a masy zón (**Zone Mass Field** – numerické). Volitelně lze zadat v dalších polích i další nastavení.

Lze vynutit tvorbu vícejaderných regionů hodnotou v poli **Zone Cooperation/Agglomeration Field**. V takovém případě je nutné zvolit jednu zónu jádra, jejíž identifikátor zadáme do tohoto pole ostatním zónám stejného jádra. Takto je možné na základě expertní znalosti specifikovat skupiny zón, které funkčně jednoznačně tvoří jeden celek – typicky pokud jejich sídla tvoří jednu souvisle zastavěnou oblast a probíhá mezi nimi územní dělba práce, jako v případě průmyslového areálu v obci těsně sousedící s městem (např. Mladá Boleslav – Kosmonosy, Beroun – Králův Dvůr).

Pro barevnou vizualizaci příslušnosti k regionům je nutné do pole **Zone Color Field** zóně zadat šestimístný hexadecimální RGB kód barvy, kterou převezme region, jehož je zóna jádrem. Tyto kódy lze zvoleným zónám vygenerovat do atributové tabulky nástrojem **Assign Colors** (více viz část 4.3.3).

Do tabulky zón směřuje i výstup – vymezení regionů je uloženo do pole **Region ID Field** ve formě identifikátoru jedné ze zón jádra regionu, k němuž zóna patří. Pokud takové pole v tabulce zón již existuje, jeho hodnoty jsou načteny a použity jako výchozí členění, jež poté algoritmus pouze ověřuje a koriguje¹². Lze takto specifikovat členění pouze zčásti a tak např. kompenzovat proběhlé územní změny nebo lokální chyby v interakčních datech.

- Tabulku interakcí (**Interactions**), na základě nichž se má vymezení provést (např. data o dojížděcí). Ta musí obsahovat pole s identifikátory zdrojové a cílové zóny (**Interaction Origin ID Field**, **Interaction Target ID Field**) a silou interakce, např. počet dojíždějících lidí (**Interaction Strength Field**).
- Tabulku sousedství zón (**Zone Neighbourhood Table**). Ta musí obsahovat dvě pole **ID_FROM** a **ID_TO**, určující jednosměrný vztah sousedství pomocí identifikátorů zón. Automatizovaná tvorba této tabulky je popsána níže v části 4.3.1.

Data o geometrii zón tento nástroj přímo nenačítá, neboť prostorové vztahy jsou vyjádřeny v tabulce interakcí a sousedství; nahrávání geometrie by tak běh nástroje pouze zpomalovalo.

Parametry vymežovací metody (pro podrobnější popis viz část 3.7) se zadávají do následujících polí:

- **Allowed Cores** – volitelná možnost zadat selekční kritérium jako *where* podmínku pro selekci nad tabulkou zón. Pokud je podmínka splněna, příslušné zóně je v prvním kroku nástroje vytvořen region; pokud ne, je na tuto zónu hned po vytvoření regionů aplikováno agregační kritérium.
- **Minimum Protoregion Hinterland EMW** – hranice EMW pro úvodní ověřovací kritérium W_0 v jednotkách masy zón.
- **Minimum Region EMW** – volitelný parametr umožňující přesně specifikovat hranici W_1 pro hlavní ověřovací kritérium a tak potlačit její výpočet pomocí generalizovaného modu.
- **Exclave Assignment Weight Percent** – stanovuje míru, s níž jsou tolerovány exklávy (koeficient penalizace exkláv k_{e0} v procentech): 0 % (vůbec) až 100 % (exklávy se nerozlišují).
- **Minimum Zone Reassignment Gain** – minimální zisk pro přerazovací kritérium r_0 (v procentech).
- **Region Overlap Merge Percent** – minimální překryv regionů pro sloučení RO_0 (v procentech).

¹²Každá zóna, jejíž pole **Region ID Field** obsahuje ne-NULL hodnotu (identifikátor), je hned na začátku běhu nástroje přiřazena jako zázemí k regionu s tímto identifikátorem.

Možnosti výstupu lze nepovinně rozšířit následovně:

- Různé charakteristiky zón a regionů (např. členské funkce, souhrnné charakteristiky interakcí) nástroj spočítá a uloží do polí v tabulce zón, pokud jsou zaškrtnuty jejich kolonky v parametru `Output Measures`. Pro více podrobností o charakteristikách viz dokumentaci nástroje.
- Informace o vícejaderných regionech vytvořených nástrojem se uchovají, pokud je zaškrtnuto políčko `Output Region Core Cooperation Field`. V tom případě je do tabulky zón zapsáno další pole, do něž se pro každou zónu tvoří jádro regionu s jiným identifikátorem, než má sama, zapíše tento identifikátor. Toto pole se následně dá využít jako `Zone Cooperation/Agglomeration Field`.
- Pro barevnou vizualizaci příslušnosti k regionům je nutné do parametru `Output Color Field` zadat název pole, do něž bude každé zóně zapsán šestimístný hexadecimální RGB kód výsledné barvy. Tyto barvy lze následně do podoby rastru vizualizovat nástrojem `Create Color Map` (viz část 4.3.3).

4.2.3 Zajištění vstupu a výstupu

Ihned po spuštění nástroje jsou nahrány jeho parametry, na jejichž základě se načtou vstupní tabulky pomocí instancí potomků třídy `Loader`¹³, využívajících `arcpy.SearchCursor`. `ZoneLoader` nahraje z tabulky zón slovník zón podle identifikátorů (`Zone`, více viz část 4.2.4, Selekční kritérium).

Následně jim `InteractionLoader` přiřadí pomocí metod `Zone.addOutflow()` a `Zone.addInflow()` interakce z nahrané tabulky (jako nové záznamy do objektů `Interactions` uložených v `Zone.outflows` a `Zone.inflows`). Pokud u daného záznamu interakce pouze jeden z obou identifikátorů odpovídá existující zóně (např. jde o záznam označující neidentifikovaný cíl vyjížďky nebo je cíl v jiném státě), je jí přiřazen jako neidentifikovaný (*raw*) tok (`Zone.addRawOutflow()` ap.)

`NeighbourLoader` nakonec pomocí `Zone.addNeighbour()` přidá zónám odkazy na jejich sousedy do seznamu `Zone.neighbours`.

Nahrávání tabulek přes `arcpy.SearchCursor` je poměrně pomalé, zvláště u interakcí, jichž je zpravidla nejvíce; po přeřazovacím kritériu jde o časově nejnáročnější část nástroje, ovšem těžko ji lze v rámci programu dále optimalizovat.

Hlavní běh nástroje je svěřen *regionalizéru* – instanci třídy `Regionaliser`¹⁴. Ten dostane slovník zón a vymezení funkčních regionů metodou popsanou v kapitole 3. Zóny jsou potom předány zpět instanci `ZoneLoader`, která pomocí `arcpy.UpdateCursor` zapíše v metodě `output()` každé zóně do pole `Region ID Field` informaci o jejím přiřazení k určitému regionu. Tato informace je představována identifikátorem jedné ze zón jádra regionu, k němuž je zóna přiřazena. To samo o sobě umožňuje regiony rychle vizualizovat; polygony regionů (tedy prostorová data) lze získat následnou aplikací nástroje `Dissolve Regions`.

¹³Ty jsou *singleton* – vždy od nich existuje pouze jedna instance.

¹⁴Také *singleton*.

4.2.4 Implementace kritérií vymezení

Hlavní běh nástroje odpovídá diagramu na obr. 8, kde jsou vyznačeny i metody regionalizéru provádějící jednotlivé části programu. Níže jsou potom detailně rozvedena jednotlivá použitá kritéria.

Selekční kritérium je provedeno rovnou při nahrávání tabulky zón. Ze zón, které jsou vybrány jako potenciální jádra, jsou vytvořeny instance třídy `CoreZone`, z ostatních `HinterlandZone`. Pokud není selekční kritérium zadáno (parametr `Allowed Cores`), jsou všechny zóny instance `CoreZone`.

Při tvorbě regionů je pro každou `CoreZone` vytvořen v metodě `createRegions()` nový `Region` s touto zónou jako jádrem (přiřazení typu `CoreAssignment`). Následně jsou v metodě `bindHinterlands()` všechny `HinterlandZones` přiřazeny pomocí agregačního kritéria k těmto regionům jako zázemí (přiřazení typu `HinterlandAssignment`).

Agregační kritérium zajišťuje třída `Bindings`, která převezme agregovaný `object` (zónu nebo region) a jeho interakce (`flows`). Jejich transformaci do agregačního kritéria zajišťuje metoda `transformFlows()`, která je volána metodou `calculate()`. Ta je provedena po každé, když se změní vymezení regionů. Objekt je poté přiřazen k takovému regionu, pro nějž je hodnota agregačního kritéria největší.

Metoda `transformFlows()` nejprve spočítá agregační kritéria podle toků do zón jako instanci `Interactions`. Následně místo cílů, jež jsou jádry regionů, dosadí do vzniklého slovníku jako klíče jejich regiony (přičemž sečte hodnoty kritérií pro vícejaderné regiony), cíle ležící v zázemí nebo doposud nepřijížené ponechá. Výsledné cíle potom seřadí do seznamu cílů (`sortedTargets`) dle hodnoty agregačního kritéria.

To umožní funkci `getMaxBind()` rychle vyhledat nejsilnější vazbu. Pokud je na prvním místě v seznamu cílů nalezena zóna, znamená to, že leží v zázemí a není to platný cíl pro agregaci – proto se provede metoda `subBind()`, která transformuje původní `flows` podle pravidla nepřímé vazby (`subBindFlows()`), posune začátek hledání v seznamu cílů o jednu pozici dál a celý proces se vrací na začátek – tak dlouho, dokud není nalezen region. K němu je vytvořeno odpovídající přiřazení (instance třídy `Assignment`) a tím jsou zóna a region svázaný (`Assignment.tangle()`).

Přiřazování rozpouštěného regionu řídí metoda regionalizéru `assignRegion()`. Ta nejprve přiřadí jádro regionu (`reassignRegionByBinds()`) a následně všechny zóny zázemí od největší po nejmenší dle masy (`reassignZones()`).

Pokud je agregační kritérium omezeno podmínkou sousedství (metodě je předáno `True` jako nepovinný *boolean* argument `neigh`), výše uvedené metody se chovají odlišně – pro zóny zázemí je zavedena prioritní fronta pomocí `Binding.level`. Nejprve jsou přiřazeny všechny zóny, pro něž primární tok (po aplikaci nepřímých vazeb) směřuje do sousedního regionu, poté ty, pro něž je nutné použít sekundární tok, poté ty, pro něž do sousedního regionu směřuje až třetí nejsilnější tok atd. Pokud jsou vyčerpány všechny toky, je zóna přiřazena dle sousedství (viz níže).

Ve vstupních datech se občas vyskytují i řídké případy, kdy zóna nemá žádnou odchozí interakci (resp. žádnou odchozí interakci do sousedního regionu) – jde např. o velmi malé obce. V tom případě je přiřazení odloženo až za všechna ostatní a provádí se na základě

sousedství – zóna je přiřazena k tomu regionu, k němuž patří nejvíce jejích sousedů (v případě rovnosti k regionu s nejvyšší EMW). Toto je považováno za nestandardní případ a na výstup je vypsáno varování (`arcpy.AddWarning()`).

Ověřovací kritérium zjišťuje regionalizér výsledkem funkce `verifyMass()` nad konkrétním regionem, jehož hodnota se porovná s aktuální hodnotou `verifyThreshold`, nastavitelnou funkcí `setVerify()`. Funkce `verifyMass()` vrací dle aktuálního kritéria buď EMW předaného regionu, nebo EMW jeho zázemí.

Generalizovaný modus EMW regionů se počítá z EMW regionů zbylých po první agregaci metodou `generalisedMode()`. Ta rozdělí soubor regionů na určitý počet kvantilů podle EMW, z nich vybere ten s nejužším rozpětím hodnot a z nich spočítá průměr. Pro větší stabilitu je tento postup opakován pro různý počet kvantilů od 6 do 20 (meze byly zvoleny na základě empirického testování) a výsledky ze všech opakování jsou zprůměrovány.

Region svoji EMW zjišťuje z aktuálních členských funkcí svých zón pomocí metody `Assignment.getDegree()`. Pro vyšší výkon se EMW cachuje (`_mass`); její opakované výpočty při každém ověření by běh programu velmi zpomalovaly, a tak se pouze po každé změně vymezení regionu aktualizuje metodou `hamplUpdateAssignments()`, kterou volá regionalizér v `updateAssignments()`. Ta nechá všechna přiřazení (`Assignment`) přepočítat hodnoty členské funkce.

Pokud je region vůči interakcím zcela uzavřený (nevede z něj žádná ven) a přitom nesplňuje ověřovací kritérium, není agregován, není totiž jak o agregaci rozhodnout. Místo toho je takový region zachován pomocí lokální výjimky uložené v `Region.indepOverride`. Pokud na něj narazí `verifyMass()`, vrátí hodnotu `verifyThreshold`.

Přeřazovací kritérium provádí regionalizér metodou `change()`, jíž jsou předány regiony, k nimž se má přeřazovat; v rámci agregačního cyklu jí `tryAddNeighbours()` předá region, jehož masu se snaží zvýšit, v optimalizační fázi přijdou od `optimize()` všechny regiony.

Vlastní přeřazení (změnu) ztělesňuje instance třídy `NeighbourChange`, obsahující odkaz na přeřazovanou zónu, zdrojový a cílový region. Při svém vytvoření spočítá zisk ze sebe plynoucí (`calcGain()`), volá `Region.getEMWDiff()`, pro výpočet viz část 3.4 na s. 30). Pokud tento zisk přesahuje stanovenou minimální hranici r_0 a zóna již nebyla do tohoto regionu přeřazována¹⁵ (`isChangeViable()`), je zařazena do seznamu změn.

V rámci výpočtu zisku plynoucího ze změny je nutné posoudit i dopad na souvislost obou regionů, přičemž tato otázka není jednoduchá a řešení je i výpočetně náročné. Exklávy mohou zaniknout přeřazením do jiného regionu, s nímž sousedí, ale také přeřazením vhodné zóny do jejich regionu, která je propojí s jeho souvislou částí. Naopak k vzniku exklávy může dojít při přeřazení zóny, která je jako jediná propojovala s jádrem regionu.

Na první z těchto tří případů pamatuje členská funkce, která se po přeřazení exklávy do souvislého regionu zvýší, protože už není násobena koeficientem penalizace. Zbylé dva případy ošetřuje metoda `Assignment.isOnlyConnection()`, volající metodu svého regionu `getArticulation()`. V případě, kdy je zóna k regionu připojována (a tedy sousedí s jeho

¹⁵Bez této podmínky může program skončit v nekonečném cyklu, kdy se stále znovu provádí tři nebo více ziskových změn.

souvislou částí), projde region své exklávy a zjistí, zda nějaká s touto zónou sousedí; pokud ano, přidá zvýšení jejich členské funkce k zisku změny.

Třetí případ, kdy je zóna z regionu odebírána, je nejsložitější – obnáší zjistit, zda odtržení zóny roztrhne region na dvě navzájem nesousedící části, vytvoří se tedy exklávy. Pokud ano, je tato zóna nazývána *artikulací* (termín z teorie grafů). Artikulace lze v regionu detekovat algoritmem založeným na procházení grafu sousedství do hloubky (*depth-first search*, *DFS*). Podobu algoritmu popisuje Černý (2010) a implementace se nachází v metodě `calcArticulations()`. Založen je na myšlence, že zóna je artikulací, pokud ze skupiny zón objevených při procházení hlouběji od ní nevede žádný vztah sousedství mimo tuto skupinu.

Metoda `change()` nejdříve vytvoří všechny změny, které přicházejí v úvahu (`getAllChanges()`). Další zpracování probíhá v cyklu. V každém opakování vybere aktuálně nejziskovější změnu a tu provede (`NeighbourChange.makeAssignment()`; seznam změn je udržován seřazený podle zisku). Pokud platí podmínka, že zdrojový region nesmí přestat splňovat ověřovací kritérium (při optimalizaci), je ověřeno a v případě neúspěchu se změna vrátí zpět (`NeighbourChange.makeRollbackAssignment()`). Jinak je změna zaregistrována (pro budoucí porovnání s následujícím přeřazením) a v seznamu změn jsou odstraněny a znovu vytvořeny všechny změny dotýkající se jejího zdrojového nebo cílového regionu (`getAffectedChanges()`)¹⁶. Změny jsou generovány metodami `getChangesToRegion()` a `getChangesFromRegion()`, jež je sestavují nad hraničními zónami svých, resp. okolních regionů.

Implementace přeřazovacího kritéria patřila v rámci nástroje k nejsložitějším, což odráží složitost kritéria samotného – v rámci každé změny je nutné posoudit dopad na členskou funkci každé zóny obou dotčených regionů, včetně souvislosti, přičemž takových změn je velké množství a po každém přeřazení se musí posuzovat znovu. Z toho vyplývá i fakt, že jde o výpočetně nejnáročnější část celého nástroje – jeho zpracování zabírá přes 90 % běhu.

Slučovací kritérium je implementováno přímočaře pomocí metody `tryMerge()`, jíž se předá region, který se snažíme sloučit s jeho sousedy. Pro každého souseda se zjistí vzájemný překryv RO , podle něj se vybere soused s největším překryvem a je-li tento větší než minimální hranice RO_0 , jsou oba regiony sloučeny – všechny zóny původního regionu jsou přiřazeny stejným typem přiřazení (`CoreAssignment` nebo `HinterlandAssignment`) k sousedovi.

4.2.5 Běh nástroje

Nástroj s výše uvedeným nastavením běží na procesoru Intel Core i7 (4 jádra o 2,1 GHz) s 8 GB RAM přibližně 15 minut, přičemž největší zátěž představuje přeřazovací kritérium, které se nepodařilo ani přes velkou snahu výrazněji zefektivnit. Zanedbatelný není ani čas nutný pro čtení a zápis dat; potřebná doba rovněž výrazně roste, pokud je zadáno velké množství výstupních charakteristik (`OutputMeasures`).

Vzhledem k nepříliš oslnivému výkonu nástroje byl tento i ostatní navrhován především s ohledem na časovou, nikoli prostorovou složitost, a některé jeho části tak nejsou příliš paměťově efektivní.

¹⁶Místo prostého přepočítání zisků všech ovlivněných změn se volí tento náročnější postup, který umožňuje zachytit i nově vzniklá sousedství.

4.3 Implementace ostatních prvků

V této části je popsána implementace ostatních prvků toolboxu, které plní převážně podpůrnou funkci.

4.3.1 Sousedství

Matice sousedství se podobně jako matice interakcí reprezentuje tabulkou obsahující dvojice identifikátorů sousedících zón, přičemž jeden řádek znázorňuje jeden směr vztahu. Jelikož je sousedství většinou symetrický vztah, bude počet řádků tabulky sudý.

Administrativní sousedství se získá prostou analýzou geometrie polygonů zón. Jde o výpočetně náročnou operaci, neboť počet nutných porovnání geometrie roste kvadraticky s počtem zón. V ArcGISu existuje několik cest, jak tabulku administrativního sousedství obdržet:

- Standardní nástroj **Near** s parametrem **Search Distance** rovným 0. Tato možnost je nejpomalejší, neboť nástroj počítá vzdálenosti, což je výpočetně náročnější operace.
- Vlastní analýza geometrie, která se nahraje přes **SearchCursor** a nad výslednými objekty se volá funkce **disjoint()**. Tento postup je implementován ve skriptu **Administrative Neighbourhood Table**. Běží zhruba dvakrát rychleji než nástroj **Near**.
- V ArcGIS modulu **Spatial Statistics** je obsažen nástroj **Generate Spatial Weights Matrix**, jehož lze po nastavení vhodných parametrů a konverzi matice prostorových vah do tabulky (**Convert Spatial Weights Matrix to Table**) využít. Jeho výhodou je řádově vyšší rychlost než u předchozích přístupů. Pro rychlé generování tabulky je postup automatizován ve skriptu **Spatial Matrix Neighbourhood Table**, kde je možné rozhodnout, zda za sousedství bude považováno i sousedství přes roh – jedním bodem (volba **CONTIGUITY_EDGES_CORNERS**, tzv. *případ královna* – viz Spurná 2008), nebo pouze hranou (*případ věž*; **CONTIGUITY_EDGES_ONLY**).

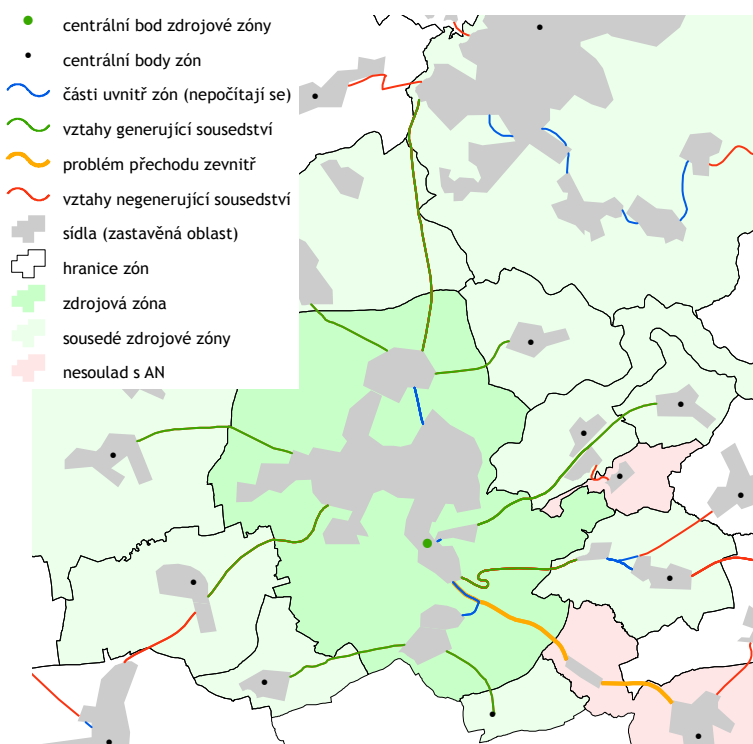
Pro urychlení běhu skriptů je dobré použít *prostorový index* (**Add Spatial Index**), který výrazně urychluje prohledávání vstupní vrstvy tím, že její prvky katalogizuje podle čtvercové sítě o třech různých rozměrech buňky.

Thiessenovy polygony kolem centrálních bodů zón lze jednoduše vygenerovat standardním nástrojem **Create Thiessen Polygons** a na ty následně aplikovat jednu z výše uvedených metod. Výsledky jsou rozporuplné; je sice odstraněn problém s nesouvislými zónami, avšak některé nesouvislosti znovu vznikají, pokud se pro vizualizaci regionů použijí administrativní hranice, a to zejména v okolí rozlohou velkých zón.

Dopravní sousedství použité v návrhu vlastní metody je počítáno skriptem **Network Neighbourhood Table**, který využívá rozšíření ArcGISu pro síťovou analýzu, **Network Analyst** (konkrétně jeho analýzu typu **Closest Facility**, která počítá cesty z daného bodu do několika nejbližších bodů sousedních, implementovanou v nástroji **Create Network Connections**).

Vstupem do nástroje jsou kromě polygonů zón i jejich centrální body (ty je možno předem spočítat jako centroidy nástrojem **Feature to Point**) a polygony sídel, zahrnující vlastní zastavěnou oblast sídel tvořících zóny. Ty by neměly přesahovat hranice polygonů zón, jinak jsou považovány za součást zóny, do níž přesahují.

Obrázek 10: Ukázka tvorby vztahů dopravního sousedství v okolí Nového Města nad Metují. Na jihovýchodním okraji je vidět příklad problému přechodu zevnitř; červenou barvou jsou označeny zóny, s nimiž není navázán vztah dopravního sousedství navzdory sousedství administrativnímu.



Zdroj: ArcČR (2013), CEDA (2010), vlastní zpracování

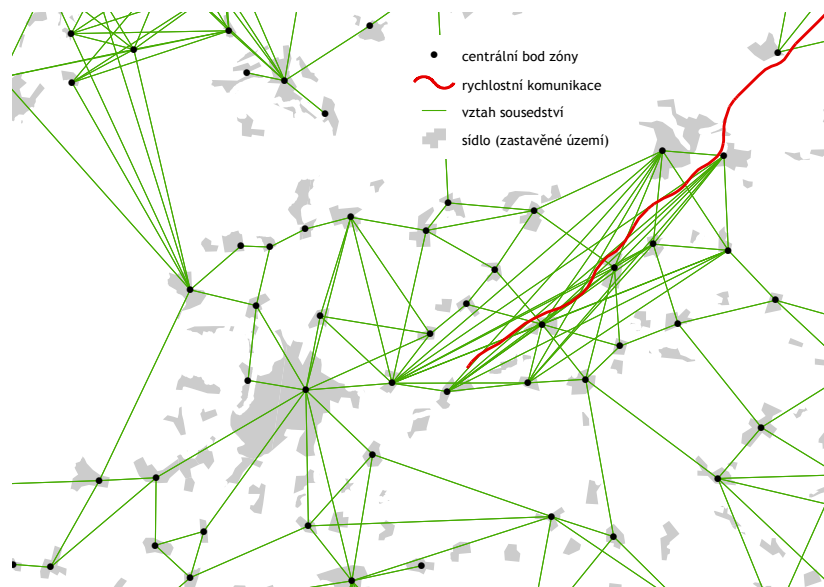
Pro každý centrální bod je nejprve nalezen určený počet linií cest k nejbližším dalším centrálním bodům. Od těchto cest jsou odečteny segmenty uvnitř polygonů sídel (**Erase**) a výsledné multipart linie jsou převedeny na singlepart (**Multipart to Singlepart**). Dále jsou vyňaty ty linie, které nepřekračují hranice polygonů zón (spojují dvě sídla v rámci jedné zóny), pomocí **Select by Location** (volba **COMPLETELY_WITHIN**). Pokud je potom ve výsledné vrstvě obsažena pouze jedna linie s danými identifikátory startu a cíle, jde o spojnici dvou zón neprocházející třetím sídlem, start sousedí s cílem a jejich (jednosměrný) vztah je zaznamenán do výstupní tabulky sousedství.

Nástroj **Network Neighbourhood Table** je velmi pomalý, zejména kvůli náročné síťové analýze a následujícím dvěma operacím překrytí. Pro tisíce zón se doba běhu počítá v hodinách. Metoda navíc trpí *problémem přechodu zevnitř* (viz obr. 10) – pokud komunikace mezi dvěma sídly téže zóny překračuje její administrativní hranici, není eliminována operátorem **COMPLETELY_WITHIN** a tudíž není s takovou zónou z daného směru navázáno sousedství¹⁷.

¹⁷To v některých případech vede i k tvorbě zón bez souseda, což je pro vymezovací algoritmus krajně nežádoucí.

Tento problém by bylo možné eliminovat převodem dílčích singlepart linií na koncové body a testováním, zda oba koncové body leží ve stejné zóně. Tento úkol je však složitější a v této práci nebyl zpracován; navíc by s sebou nesl další významný nárůst výpočetní náročnosti.

Obrázek 11: Ukázka výsledných dopravně-sousedských vztahů v okolí Příbrami. Přítomnost rychlostní silnice R4 sousedské vztahy ztlačuje.



Zdroj: ArcČR (2013), CEDA (2010), vlastní zpracování

Další problém při použití nástroje souvisí se selekcí vhodné sítě. Při aplikaci na silniční síť se ukázalo, že do sítě pro kalkulaci sousedství není vhodné zahrnovat dálnice a některé silnice vyšších tříd, neboť ty se z velké části vyhýbají zastavěným oblastem a umožňují tak bez průchodu jiným sídlem propojit řadu i velmi vzdálených zón (viz obr. 11).

4.3.2 Vzdálenost

Vzdálenost, přes níž interakce působí, se získá převedením tabulky interakcí na liniovou *feature class*. Ta, pokud je uložena v geodatabázi, automaticky získá atribut *Shape_Length*, jehož lze použít pro další analýzy.

Převod interakcí na liniovou vrstvu zajišťují dva skripty v toolsetu *Conversion – Table to Interactions*, který tvoří přímé čáry mezi startem a cílem, tedy vzdušnou vzdálenost, a *Table to Connections*, který tvoří spojnice (a tedy i vzdálenosti) po zadané dopravní síti. Oba skripty vyžadují jako vstup bodovou vrstvu centrálních bodů zón, *Table to Connections* navíc i vybudovaný síťový dataset (*network dataset*; je tedy nutné mít licenci na Network Analyst). Tvorba síťových spojení nástrojem *Table to Connections* je samozřejmě podstatně pomalejší než tvorba přímých čar, které vytváří *Table to Interactions*; rozdíl v rychlosti tvoří zhruba dva řády a pro desetitisíce řádků se začíná doba běhu nástroje *Table to Connections* počítat v hodinách. Výhodou je však výsledek reflektující realitu daleko přesněji.

Pro prostou eliminaci interakcí působících na příliš dlouhou vzdálenost stačí jednoduché atributové početní operace. Ty jsou implementovány ve skriptu *Select Interactions*,

kde je možné vyloučit interakce příliš slabé, příliš dlouhé nebo s příliš malou relativní silou $\left(\frac{f_{ij}}{d_{ij}}\right)$. Sofistikovanější metodou je aplikace *distance-decay* funkce. Jedna z takových možností, využívající Gaussovu křivku, je popsána v části 3.2 na s. 25 a implementována ve skriptu `Gaussian Distance Decay`.

4.3.3 Vizualizace příslušnosti zón k regionům

Vizualizace příslušnosti zón k regionům dle postupu popsaného v části 3.8 na s. 32 je zajišťována dvěma nástroji.

Nástroj `Assign Colors` přiřazuje vstupním zónám podle předložené tabulky sousedství syté barvy, které uloží do jejich atributové tabulky jako hexadecimální RGB kód (např. `ff0000` pro červenou). Tato reprezentace je výhodná svou rozšířeností, kompaktností a zároveň snadným převodem do číselných RGB hodnot; navíc je používána v HTML, takže případná tvorba on-line mapy může být relativně snadná. Výchozí syté barvy jsou předem určeny v souboru `color_bright.conf` v adresáři s toolboxem a je tak možné je rozšířit nebo předefinovat.

Při přiřazování barev zónám bylo použito heuristiky, již uvádí Večerka (2007, s. 34). Ta sleduje minimalizaci množství použitých barev při zachování podmínky, že sousedé nesmí být obarveni stejně, a to tak, že nejprve obarvuje nejvíce zapojené vrcholy (řazení nejprve podle počtu obarvených a sekundárně podle počtu neobarvených sousedů). Barvy drží uspořádané a každou zónu se snaží obarvit barvou co nejbližší začátku jejich seznamu.

Minimalizace počtu barev však není hlavním cílem vizualizace příslušnosti – naopak, je vhodné, aby blízké, byť nesousedící regiony byly obarveny různě, neboť jinak by nešlo rozlišit jejich sféry vlivu v regionech, které je oddělují. Proto je v použité implementaci seznam barev před obarvením každé zóny náhodně přeuspořádán a teprve následně je vybrána (ve shodě s původní metodou) první barva, již není obarven žádný ze sousedů.

Nástroj `Delimit Functional Regions` poté spočítá pro každou zónu její členské funkce (pokud jsou zadány příslušné parametry) a uloží do atributové tabulky RGB kód odpovídající barvě smíšené podle příslušnosti k regionům.

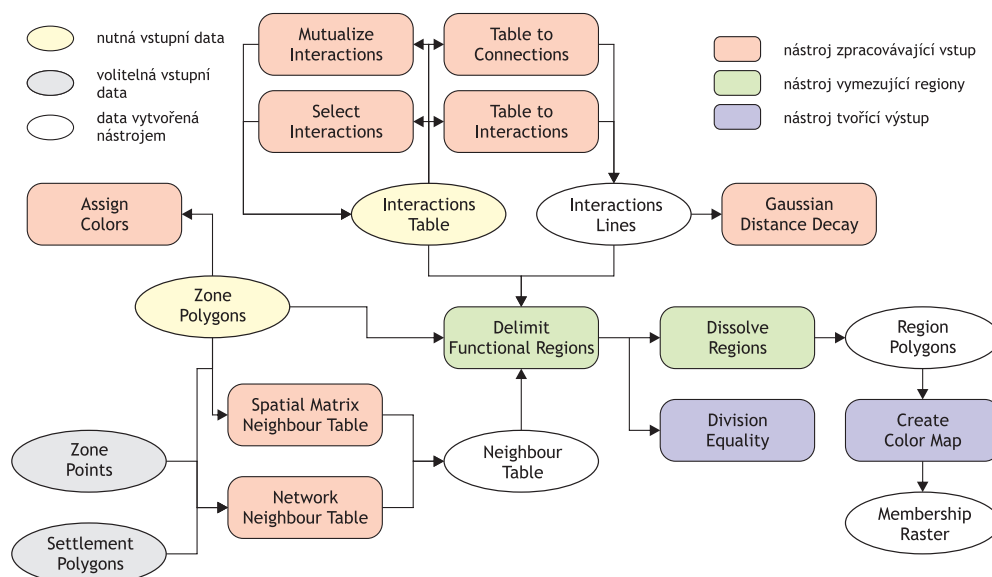
Protože ArcGIS zatím neumí při vizualizaci přístupným způsobem využít přesnou barvu definovanou v atributové tabulce, nabízí se pouze převod na barevný ras. To zajišťuje nástroj `Create Color Map`, který ve vstupní polygonové vrstvě zón nejprve vytvoří tři pole odpovídající RGB složkám s číselným údajem intenzity, následně třikrát spustí nástroj `ArcGISu Raster to Polygon` nad těmito údaji, čímž vytvoří tři černobílé rastry, které v závěru zkombinuje do jednoho barevného (`Composite Bands`).

Protože některé jiné formáty než ESRI Grid neumí správně rozlišit `NoData` hodnoty od nulových a při jejich použití se objevovaly ve výstupu „díry“ tam, kde byla jedna z RGB složek nulová, probíhají všechny výpočty ve formátech ESRI Grid a teprve výsledek je konvertován do zadaného formátu (`Copy Raster`) – je doporučován TIFF, neboť jej lze snadno georeferencovat.

4.4 Postup použití

Pro vymezování funkčních regionů je nutné si nejprve připravit vstupní data – polygonovou *feature class* zón a tabulku interakcí.

Obrázek 12: Schéma předpokládaného použití toolboxu. Pro přehlednost obsahuje pouze vstupy tabulek a geometrických vrstev, nikoli číselných a textových parametrů.



Tabulku lze specifikovat v různých formátech – většinu běžných formátů včetně Microsoft Excelu ArcGIS bez problémů přečte –, ovšem je vhodné ji před použitím exportovat do geodatabáze (výhodou je vyšší rychlost a možnost zápisu).

Je nezbytné, aby si tyto zóny a interakce odpovídaly jak časově (vzhledem k případným územním změnám), tak identifikátory – zóny musí v atributové tabulce obsahovat shodný typ identifikátoru, jaký používají interakce pro označení výchozí a cílové zóny. Nejvhodnější je použít standardní číselníkové identifikátory (v Česku IČZUJ), zásadně ne názvy, neboť jsou náchylné k chybám zejména kvůli znakům národních abeced, různým zkratkám a podobně.

Interakce lze před vstupem ještě různě modifikovat pomocí vzdálenostní korekce (**Gaussian Distance Decay**, viz část 4.3.2), vybírat, převést na symetrické a podobně.

Následně je třeba vygenerovat tabulku sousedství některým z nástrojů toolsetu **Neighbourhood**, popsaných v části 4.3.1, nejlépe také do geodatabáze, a na základě expertní znalosti určit parametry regionalizace. Vše se pak spolu s požadavky na výstupy zadá do nástroje **Delimit Functional Regions**, který vypíše údaje o vymezení regionů a případně další požadované výstupy (viz část 4.2.2) do tabulky zón.

Tyto výstupy lze rychle vizualizovat v ArcMapu (např. regionální příslušnost v **Symbology/Categories**). Nástrojem **Dissolve Regions** můžeme získat polygony vymezených regionů s atributy obsahujícími agregátní data za dané regiony.

Pokud má být vymezena hierarchická soustava regionů na více řádovostních úrovních, je nutné nástroj **Delimit Functional Regions** a **Dissolve Regions** spustit opakovaně pro každou řádovostní úroveň od nejnížší k nejvyšší. Po každé úrovni je nutné také agregovat interakce na regionální úroveň (z interakcí mezi zónami se stanou interakce v rámci regionů), pro což slouží nástroj **Aggregate Flows**.

4.5 Regionalizace Česka

Tato část obsahuje ukázkou aplikace výše popsané vlastní regionalizační metody implementované v nástroji `Delimit Functional Regions` na denní dojíždku popsanou Sčítáním lidu, domů a bytů (SLDB) 2011.

4.5.1 Použitá data a parametry

Data ze SLDB 2011 jsou dostupná na úrovni obcí, tedy základních administrativních jednotek, skládajících se mnohdy z více sídel; MAUP zkreslení je však stále zanedbatelné (viz část 2.2 na s. 10). Obce, jichž je 6251, tedy pro tuto úlohu tvoří zóny.

Konkrétně jsou použita data o pracovní a školské dojíždce (ČSÚ 2013), která jsou dostupná až do úrovně jednotlivých dojíždějících – z nich byly kvůli neúplnosti jako síly interakce převzaty údaje o celkové souhrnné denní pracovní i školské dojíždce (Hampl a Marada 2014). Tabulka čítá přibližně 180 000 interakcí a je ve formátu `.dbf` uložena na CD v příloze 1. Významy jednotlivých polí jsou uvedeny v tabulce 1 (v této práci byly použity údaje z pole `DEN`).

Tabulka 1: Vysvětlení obsahu jednotlivých polí tabulky interakcí, která je uložena na CD v příloze 1.

Název pole	Obsah
<code>O_ICZUJ</code>	identifikátor IČZUJ výchozí zóny interakce
<code>D_ICZUJ</code>	identifikátor IČZUJ cílové zóny interakce
<code>CELK</code>	celková dojíždka ($CELK = PRAC + SKOL$)
<code>PRAC</code>	celková dojíždka za práci (včetně nedenní ^a)
<code>PRACDEN</code>	denní dojíždka za práci (počet lidí dojíždějících denně)
<code>SKOL</code>	celková dojíždka do škol všech typů a stupňů (včetně nedenní)
<code>SKOLDEN</code>	denní dojíždka do škol všech typů a stupňů
<code>DEN</code>	souhrn denní dojíždky za práci i do škol ($DEN = PRACDEN + SKOLDEN$)

^aNedenní dojíždkou se rozumí dojíždka sice opakovaná, avšak ne každodenní a ne nutně pravidelná (např. pouze dva dny v týdnu). Více viz ČSÚ (2013).

Počty obyvatel obcí k datu sčítání byly spolu s prostorovým vymezením převzaty zprostředkovaně z volně dostupné databáze ArcČR (ArcČR 2013, polygonová vrstva `ObcePolygony` s ručně vrácenými územními změnami pomocí dat z vrstvy `CastiObcePolygony` do stavu k 15. 3. 2011). Jako mas bylo využito populačních velikostí z pole `OB11`.

Sousedství bylo pro finální verzi vygenerováno pomocí nástroje `Spatial Matrix Neighbour Table` s parametrem `CONTIGUITY_EDGES_ONLY`. Původní záměr použít dopravní sousedství selhal na problémech představených v části 4.3.1; bylo sice otestováno jako použitelné na dílčích územích, pro ostré nasazení však bude ještě potřebovat vylepšit.

Několik zón bylo – podle seznamu, jež uvádí ve své publikaci Hampl (2005)¹⁸ – manuálně zařazeno do vícejaderného regionu pomocí `Zone Cooperation/Agglomeration Field` (viz část 4.2.2).

Nastavení parametrů bylo použito následující:

¹⁸Oproti tomuto seznamu byla navíc přidána i dvojice Kolín-Ovčáry, reflektující vznik automobilky TPCA v průmyslové zóně na okraji Kolína, avšak již ne v jeho administrativních hranicích.

- Selekční kritérium (**Allowed Cores**) bylo pro zrychlení běhu aplikováno tak, že regiony byly vytvořeny pouze pro zóny s více než 1 000 obyvatel. Tato velikost tvoří přibližně 15 % velikosti nejmenšího regionálního jádra, jež vymezil Hampl (2005).
- Minimální úvodní EMW zázemí $W_0 = 1000$, což je stanoveno bohatě pod Hamplovu mez 5 000 ve snaze touto mezí neeliminovat žádné regiony s nezanedbatelným zázemím a pro jednoduchost stejně jako hranice selekčního kritéria (ovšem nutno mít na paměti, že zde jsou masy zón pro výpočet EMW váženy členskou funkcí).
- Minimální EMW regionu (**Minimum Protoregion Hinterland EMW**) W_1 byla určena programem pomocí generalizovaného modu.
- Koeficient penalizace exkláv $k_{e0} = 0$ pro zajištění principu souvislosti.
- Minimální zisk pro přerazovací kritérium $r_0 = 10\%$ (odhadem na základě analýzy výsledků).
- Minimální překryv regionů pro sloučení $RO_0 = 50\%$ (odhadem na základě analýzy výsledků; nahraává častější tvorbě vícejaderných regionů, aby byly odhaleny zajímavé vztahy).

Tyto hodnoty jsou zároveň v nástroji nastaveny jako výchozí (kromě selekčního kritéria, které je standardně prázdné).

4.5.2 Výsledek

Výsledné vymezení funkčních regionů¹⁹ je vizualizováno na mapě v příloze 3. Kvantitativní charakteristiky vymezených regionů jsou uvedeny v příloze 6. Na mapách na obr. 13 a 14 je možné toto vymezení porovnat s administrativním členěním Česka na úrovni ORP, jež mu řádovostně nejlépe odpovídá, a s vymezením, jež provedli Hampl a Marada (2014).

Tabulka 2: Porovnání agregátních charakteristik regionů vymezených programem se studií Hampl a Marada (2014). Program vymezil i menší regiony.

Populační velikost	Program	Hampl a Marada (2014)
Počet	179	149
Minimum	7 226	10 055
Medián	26 493	37 147
Průměr	59 007	70 041
Maximum	1 791 964	1 804 343
Směr. odch.	151 022	166 343

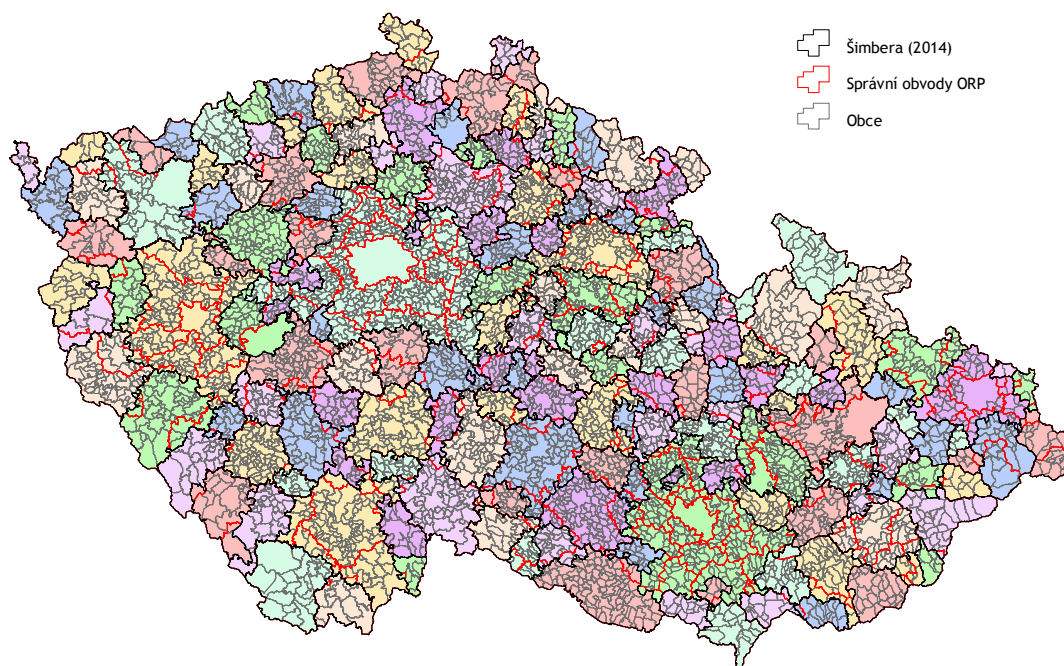
Kvantitativní porovnání s vymezením, jež provedli Hampl a Marada (2014)²⁰, je uvedeno v tabulce 2. Z ní je zřejmé, že program vymezil regiony na mírně nižší řádovostní úrovni než Hampl a Marada (2014); jejich kritérium je však založeno na analýze čtyřicet let starého sídelního systému a dodnes se zachovalo především kvůli možnosti provádět vývojová srovnání, proto lze tvrdit, že zde použitá metoda (jež mez W_1 stanovila na 2 254) věrněji odráží současný stav, byť je nutno uznat, že některé jí vymezené regiony mají velmi nízkou míru autonomie, jak je vidět na mapě v příloze 4.

Další možnosti vizualizace výsledku ukazují mapy v přílohách 4 (vizualizace příslušnosti zón k regionům dle postupu popsaného v části 4.3.3) a 5 (zobrazení hodnot členských funkcí

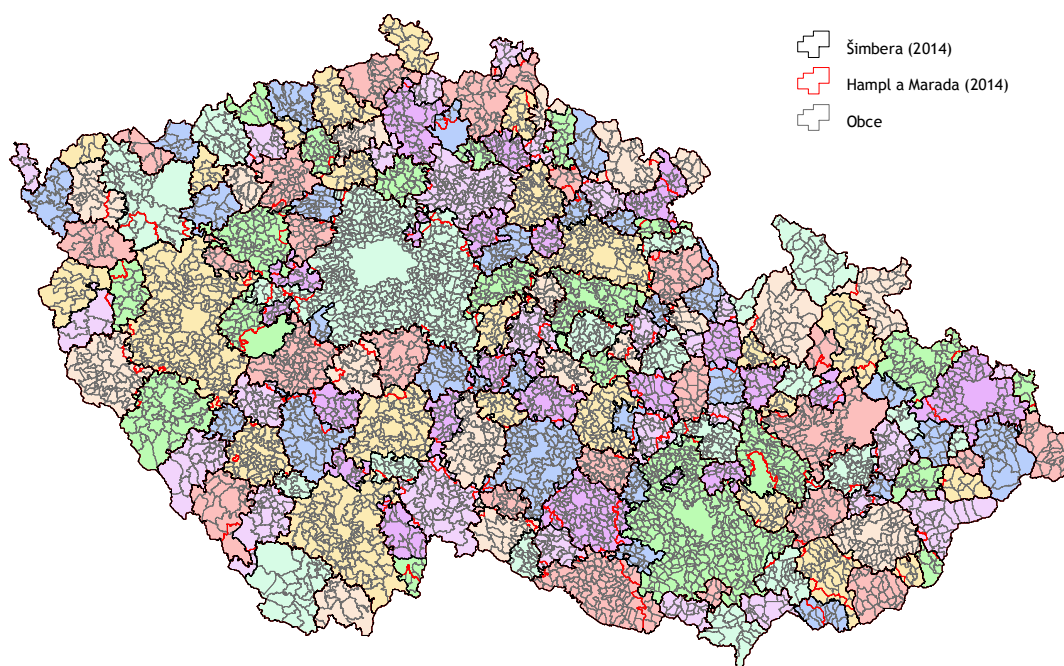
¹⁹Vzhledem k použitým interakcím by nejpřiléhavější název pro tyto regiony zněl *funkční sociogeografické mikroregiony*, neboť denní dojíždka integruje regiony relativně nízké řádovostní úrovně.

²⁰Jiné další studie zatím nad těmito daty nebyly vyhotoveny.

Obrázek 13: Přehledová mapa vymezených funkčních regionů Česka dle dat ze SLDB 2011. Oproti administrativnímu členění jsou regiony největších center výrazně větší.



Obrázek 14: Přehledová mapa vymezených funkčních regionů Česka dle dat ze SLDB 2011. Počet a umístění regionů se vcelku kryje s vymezením, jež provedli Hampl a Marada (2014).



zón v regionech, kam byly přiřazeny). V mapách jsou velmi dobře patrné anomálie tvořené vojenskými újezdy, jež se svou velkou plochou a velmi nízkou hustotou zalidnění z českého sídelního systému vymykají.

5 Diskuze výsledků

Použitá metoda regionalizace dosahuje vysoké shody s jinými v Česku použitými metodami – rozdíly se týkají především malých venkovských obcí v oscilačních zónách a regionů s nízkou integritou, které zde použitá metoda nevymezí kvůli vysoké míře vlivu uzavřenosti na ověřovací kritérium (skrz členské funkce). S úspěchem jsou detekovány některé dvojjaderné regiony (Žamberk-Letohrad) vymezené M. Hamplem; v okolí autora bydliště výsledek lépe odpovídá autorově lokální znalosti. Metoda však vymezuje (patrně díky absenci lidského faktoru) regiony s velmi členitými hranicemi, které vypadají místy velmi nerealisticky; nabízelo by se přidat pomocné přerazovací kritérium podle délky hranice.

Metoda je citlivá na neúplnost datového souboru, není tak vhodné vybírat ke zpracování pouze část interakcí s dostatečnou absolutní silou, neboť pak je značně ovlivněna členská funkce jádrových zón. Jako úspěšnou lze hodnotit implementaci výpočtu generalizovaného modu, jehož hodnotu program určil na $W_1 = 2254$. Tímto i ostatními parametry metoda sleduje minimalizaci množství v sobě zabudovaných předpokladů a parametrů; místo toho, aby bylo nutno parametry kalibrovat na daný sídelní systém, provádí kalibraci program sám a zbylými parametry naopak umožňuje výzkumníkovi stanovit prioritu jednotlivých principů vymezování (např. nakolik má být princip souvislosti předřazen principu autonomie).

Metoda přímo vychází z Hamplova nodálního konceptu regionů. Ačkoliv Casado-Díaz a Coombes (2011) tento koncept kritizují jako zastaralý a neodrážející rostoucí polycentricitu, dle současných poznatků je – alespoň v českém sídelním systému – stále dominantní (Hampl 2005) a díky umožnění existence vícejaderných regionů je i poměrně flexibilní. Snahou práce však bylo i vnést do existující plejády metod nové nápady, a proto není použitá metoda přímou kopií Hamplovy, nýbrž v sobě obsahuje nové formy kritérií, jež se z určitého úhlu pohledu mohou zdát jako experiment.

Agregační kritérium preferuje menší regiony možná až přílišnou měrou, jak je zřejmé z mapy v příloze 4, kde jsou některé regiony tvořeny v podstatě pouze jádrem a oscilační zónou; na druhou stranu mohou být takto vyjádřeny v datech nepodchycené lokální toky.

Členské funkce jsou snahou vnést do posuzování velikosti a uzavřenosti nový, matematicky smysluplný pohled, byť míra vzájemné kompenzace těchto dvou ukazatelů, jež nabízejí, je oproti dosavadním metodám výrazně větší. Koeficient penalizace exkláv pak umožňuje daleko flexibilnější práci s nesouvislými regiony.

Naopak problematické zůstává slučovací kritérium, u něž je složité stanovit mezní míru překryvu RO_0 ; podobně i přerazovací kritérium, jež je na svůj přínos poměrně složité a přesto lze některá přerazení provedená ve snaze optimalizovat výsledek s úspěchem rozporovat.

Celková programová konstrukce by si zasloužila vylepšit; přestože nástroje fungují, s dostatečným odstupem by mohla být nalezena i vhodnější struktura tříd a kvalitnější imple-

mentace zejména vzhledem k čistotě kódu a snadnosti dalších úprav.

Použití toolboxu je relativně rychlé a přímočaré – pokud jsou připravena data v tabulkách, stačí je pouze spouštět nad daty a postupně vznikajícími výstupy a po skončení běhu výsledek vizualizovat zobrazením dle kategorického atributu (pokud ArcGIS nevolí kvalitativní barvy vhodně, např. přiřazuje vedle sebe ležícím regionům velmi podobné, lze použít nástroj **Assign Colors** a jako kategorický atribut použít vytvořený barvený kód). Po celou dobu přitom není nutná žádná konverze dat a celý proces od nahrání dat po tvorbu mapového výstupu trvá i méně zkušenému uživateli nanejvýš hodinu (byť by rychlost nástrojů samozřejmě šlo dále vylepšit).

Použitím GIS se navíc otevírají daleko širší možnosti – jednak při zpřesnění výsledku jednoduchou aplikací sousedství a vzdálenosti, jednak při vizualizaci. Vytvořený toolbox zpřístupňuje tyto možnosti v poměrně značném rozsahu a nabízí se k dalšímu rozšiřování.

Dopravní sousedství dává pro většinu území smysluplné výsledky. Trpí sice *problémy přechodu zevnitř*, v některých územích kolem dálnic rovněž není příliš úspěšné a nástroj sám má velké rezervy, co se týče výkonu, což brání v jeho rutinním používání; tyto nedostatky jsou však dle autorova názoru odstranitelné a metoda stojí za další rozvíjení, neboť výrazně realističtější model sousedství, který poskytuje, může velmi pozitivně ovlivnit výsledné členění. Zamyslet by se šlo zejména nad specifikací požadavků na vstupní síť nebo použitím *in-memory* úložiště pro zrychlení overlay operací.

Barevná vizualizace příslušnosti zón k regionům pomocí rastru je sice krkolomné řešení, obcházející neschopnost ArcGISu využít v symbologii přesnou barvu definovanou jako atribut, nicméně kartograficky jde o celkem vyhovující výstup (na poměry automatického nástroje), který může poskytnout velmi užitečný vhled do situace v regionálním systému; má ovšem spíše analytický význam, neboť generované barvy jsou z kartografického hlediska neúnosně syté.

6 Závěr

Je zřejmé, že GIS může být při vymezení funkčních regionů geografovi velmi platným pomocníkem. Díky integraci všech potřebných kroků do jednoho prostředí a vzniklé synergie vytvořených vymežovacích skriptů s dalšími nástroji GIS od editace po tvorbu map odpadá nutnost zdlouhavých a chybových konverzí, takže cesta od podkladových dat ke kartografickému nebo jinému výstupu je přímočará a rychlá. Významná výhoda tkví zejména v možnosti rychle vymezení opakovat nad odlišnými daty a porovnávat např. regiony podle pracovní a školské dojížděky či podle pohlaví. Integrace nových přístupů k uchopení sousedství umožňuje regionalizaci v některých územích značně zpřesnit. Vytvořený toolbox je snadno doplnitelný a rozšiřitelný o další nástroje i možnosti konfigurace – nabízí se např. zkoumání dalších přístupů k formalizaci sousedství, vylepšování těch stávajících, sofistikovanější zahrnutí vzdálenosti nebo integrace dalších vymežovacích metod, jejichž výsledky by tak bylo možné mezi sebou jednoduše porovnávat.

Samotná metoda vymezení je jen dalším krokem na cestě vylepšování metod jí podobných; jak podotýká Coombes a kol. (1979), žádný algoritmus nemůže zcela vystihnout komplexitu sídelního systému, zvláště funkčních regionů, které se vyznačují jednou z nejnižších měr integrity z reálných systémů vůbec (Hampl 1998). Přesto je potřebné snažit se realitu vystihnout co nejpřiléhavěji a pokračovat v sestavování jejích modelů. V tomto ohledu slibují velký přínos evoluční algoritmy, které řeší řadu problémů zde představené metody.

Použité zdroje

- ANDERSEN, A. K. (2002): *Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark?*. Regional Studies, 36, č. 8, s. 833–844.
- ArcČR 500 - Digitální geografická databáze 1 : 500 000. Verze 3.1. ARCDATA PRAHA, Praha, 2013. Dostupné z: <http://www.arcdata.cz/produkty-a-sluzby/geograficka-data/arccr-500/> (cit. 27. 4. 2014).
- ArcGIS for Desktop documentation. Verze 10.0. ESRI, 2014. Dostupné z: <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html> (cit. 19. 5. 2014).
- AUERBACH, F. (1913): *Das Gesetz der Bevölkerungskonzentration*.
- BEZÁK, A. (2000): *Funkčné mestské regióny na Slovensku*. Geographia Slovaca 15. Geografický ústav SAV, Bratislava, 90 s.
- BOND, S., COOMBES, M. (2007): *2001-based Travel-To-Work Areas Methodology*. Office for National Statistics, London, 3 s.
- BROWN, A., FERINGA, W. (2003): *Colour Basics for GIS Users*. Prentice Hall, Harlow, 172 s.
- BROWN, L. A., HOLMES, J. (1971): *The delimitation of functional regions, nodal regions, and hierarchies by functional distance approaches*. Journal of Regional Science, 11, č. 1, s. 57–72.
- BURROUGH, P. A., MCDONNELL, R. A. (1998): *Principles of Geographical Information Systems*. Oxford University Press, New York, 339 s.
- CASADO-DÍAZ, J. M. (2000): *Local Labour Market Areas in Spain: A Case Study*. Regional Studies, 34, č. 9, s. 843–856.
- CASADO-DÍAZ, J. M., COOMBES, M. (2011): *The delineation of 21st century local labour market areas: a critical review and a research agenda*. Boletín de la Asociación de Geógrafos Españoles, 57, s. 7–32.
- ČR 1 : 150 000 – digitální geografická databáze. CEDA – Central European Data Agency, Praha, 2010.
- COOMBES, M. G., DIXON, J. S., GODDARD, J. B., OPENSHAW, S., TAYLOR, P. J. (1978): *Towards a more rational consideration of census areal units: daily urban systems in Britain*. Environment and Planning A, 10, s. 1179–1185.
- COOMBES, M. G., DIXON, J. S., GODDARD, J. B. a kol. (1979): *Daily urban systems in Britain: from theory to practice*. Environment and Planning A, 11, s. 565–574.
- COOMBES, M. G., DIXON, J. S., GODDARD, J. B., OPENSHAW, S., TAYLOR, P. J. (1982): *Functional regions for the population census of Great Britain*. In: Herbert, D. T., Johnston, R. J. (eds.): *Geography and the urban environment: progress in research and applications*. Wiley, Chichester, s. 64–112.

- COOMBES, M. G., GREEN, A. E., OPENSHAW, S. (1986): *An efficient algorithm to generate official statistical reporting areas: the case of the 1984 travel-to-work areas revision in Britain*. Journal of the Operational Research Society, 37, č. 10, s. 943–953.
- COOMBES, M. G. (2000): *Geographic information systems: a challenge for statistical agencies*. Research in Official Statistics, 2000, č. 2, s. 74–88.
- CÖRVERS, F., HENSEN, M., BONGAERTS, D. (2009): *Delimitation and Coherence of Functional and Administrative Regions*. Regional Studies, 43, č. 1, s. 19–31.
- ČERNÝ, J. (2010): *Základní grafové algoritmy*. Verze 0.95. MFF UK, Praha, 204 s. Dostupné z: <http://kam.mff.cuni.cz/~kuba/ka/ka.pdf> (cit. 11. 5. 2014).
- Dojíždka do zaměstnání a škol dle Sčítání lidu, domů a bytů 2011 - Česká republika*. Český statistický úřad, Praha, 2013. Dostupné z: <http://www.czso.cz/csu/2013edicniplan.nsf/p/22000-13> (cit. 27. 4. 2014).
- DONG, P. (2008): *Weighted Voronoi Diagram Extension for ArcGIS: User's Guide*. University of North Texas. Dostupné z: <http://geography.unt.edu/~pdong/software/VoronoiHelp.htm> (cit. 16. 4. 2014).
- FABRIKANT, S. I. (2001): *Visualizing region and scale in information spaces*. Proceedings, The 20th International Cartographic Conference, ICC 2001, Beijing, China, s. 2522–2529.
- FARMER, C. J. Q., FOTHERINGHAM, A. S. (2011): *Network-based functional regions*. Environment and Planning A, 43, s. 2723–2741.
- FELDMAN, O., SIMMONDS, D., TROLL, N., TSANG, F. (2005): *Creation of a System of Functional Areas for England and Wales and for Scotland*. MVA, London, 15 s.
- FENG, Z. (2009): *Fuzziness of Travel-to-Work Areas*. Regional Studies, 43, č. 5, s. 707–720.
- FINDLAY, A., SLATER, P. B. (1981): *Functional regionalization of spatial interaction data: a comment*. Environment and Planning A, 13, s. 645–646.
- FLÓREZ-REVUELTA, F., CASADO-DÍAZ, J. M., MARTÍNEZ-BERNABEU, L. (2006): *An Evolutive Approach for the Delineation of Local Labour Markets*. In: Runarsson, T. P. a kol. (eds.): *Lecture Notes in Computer Science 4193*. Springer-Verlag, Berlin Heidelberg, s. 342–351.
- FORTUNATO, S. (2010): *Community detection in graphs*. Physics Reports, 486, s. 75–174.
- GUOQING, D. (2001): *Using GIS for analysis of urban systems*. GeoJournal, 52, s. 213–221.
- HAMPL, M., GARDAVSKÝ, V., KÜHNEL, K. (1987): *Regionální struktura a vývoj systému osídlení ČSR*. Univerzita Karlova, Praha, 255 s.
- HAMPL, M. (ed.) (1996): *Geografická organizace společnosti a transformační procesy v České republice*. Přírodovědecká fakulta UK v Praze, Praha, 396 s.
- HAMPL, M. (1998): *Realita, společnost a geografická organizace: hledání integrálního řádu*. Přírodovědecká fakulta Univerzity Karlovy v Praze a DemoArt, Praha, 110 s.
- HAMPL, M. (2005): *Geografická organizace společnosti v České republice: transformační procesy a jejich obecný kontext*. PŘF UK, DemoArt, Praha, 147 s. Dostupné z: <http://www.natur.cuni.cz/geografie/vyzkumny-zamer-geograficke-sekce/publikace/geograficka-organizace-spolecnosti-v-ceske-republice> (cit. 20. 11. 2013)
- HAMPL, M., MARADA, M. (2014): *Sociogeografická regionalizace Česka 2011*. Geografie, 44 (v recenzním řízení).

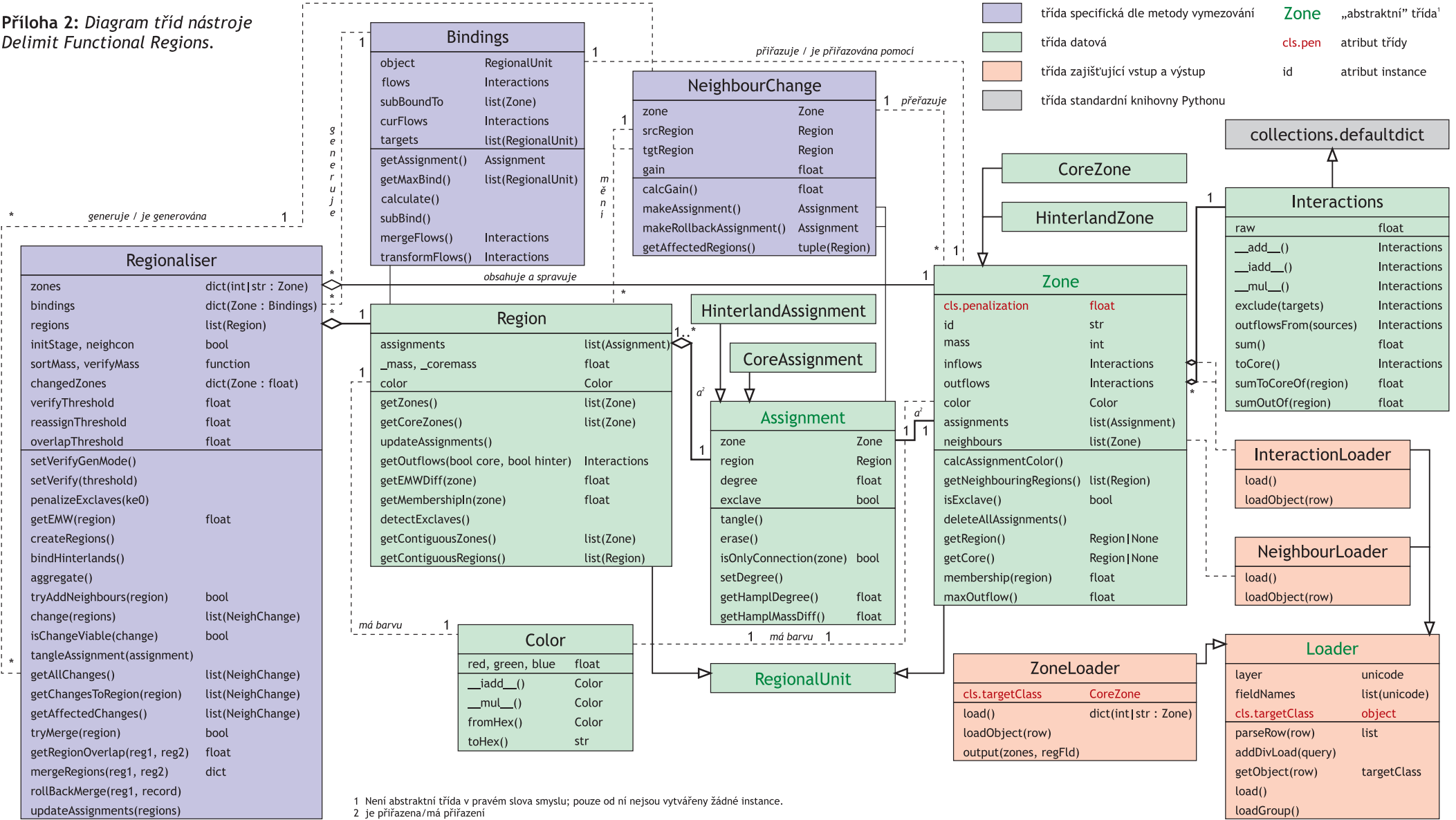
- DE JONG, T., VAN DER VAART, N. (2013): *Manual Flowmap 7.4.2*. Faculty of Geographical Sciences, Utrecht University, Utrecht, 167 s. Dostupné z: <http://flowmap.geo.uu.nl> (cit. 14. 02. 2014).
- VAN DER LAAN, L., SCHALKE, R. (2001): *Reality versus Policy: The Delineation and Testing of Local Labour Market and Spatial Policy Areas*. European Planning Studies, 9, č. 2, s. 201–221.
- LANDRÉ, M. (2012): *Geoprocessing Journey-to-Work Data: Delineating Commuting Regions in Dalarna, Sweden*. ISPRS International Journal of Geo-Information, 2012, č. I, s. 294–314.
- LANDRÉ, M., HÅKANSSON, J. (2013): *Rule versus Interaction Function: Evaluating Regional Aggregations of Commuting Flows in Sweden*. European Journal of Transport and Infrastructure Research, 13, č. 1, s. 1–19.
- KARLSSON, C., OLSSON, M. (2006): *The identification of functional regions: theory, methods, and applications*. Annals of Regional Science, 40, s. 1–18.
- KLAPKA, P., HALÁS, M., TONEV, P. a kol. (2013): *Functional regions of the Czech Republic: Comparison of simpler and more advanced methods of regional taxonomy*. Acta Universitatis Palackianae Olomucensis – Geographica, 44, č. 1, s. 45–57.
- KONJAR, M., LISEC, A., DROBNE, S. (2010): *Methods for delineation of functional regions using data on commuters*. 13th AGILE International Conference on Geographic Information Science, Guimarães, 10 s. Dostupné z: http://agile2010.dsi.uminho.pt/pen/ShortPapers_PDF/93_DOC.pdf (cit. 19. 3. 2014).
- KRAFT, S., BLAŽEK, J. (2012): *Spatial interactions and regionalisation of the Vysočina Region using the gravity models*. Acta Universitatis Palackianae Olomucensis – Geographica, 43, č. 2, s. 65–82.
- KRYGSMAN, S., DE JONG, T., NEL, J. (2009): *Functional transport regions in South Africa: An examination of national commuter data*. Proceedings of the 28th Southern African Transport Conference, s. 144–154.
- KUNC, J. a kol. (2013): *Časoprostorové modely nákupního chování v České republice*. Masarykova univerzita, Brno, 196 s.
- MARTÍNEZ-BERNABEU, L., FLÓREZ-REVUELTA, F., CASADO-DÍAZ, J. M. (2012): *Grouping genetic operators for the delineation of functional areas based on spatial interaction*. Expert Systems with Applications, 39, s. 6754–6766.
- MASSER, I., BROWN, P. J. B. (1975): *Hierarchical procedures for interaction data*. Environment and Planning A, 7, s. 509–523.
- MASSER, I., SCHEURWATER, J. (1980): *Functional regionalisation of spatial interaction data: an evaluation of some suggested strategies*. Environment and Planning A, 12, s. 1357–1382.
- MITCHELL, W., WATTS, M. (2010): *Identifying Functional Regions in Australia Using Hierarchical Aggregation Techniques*. Geographical Research, 48, č. 1, s. 24–41.
- NASCIMENTO, M. C. V., DE CARVALHO, A. C. P. L. F. (2011): *Spectral methods for graph clustering – A survey*. European Journal of Operational Research, 211, s. 221–231.
- NEWELL, J., PERRY, M. (2004): *Functional labour markets revealed by travel to work data 1991 and 2001*. Massey University, Wellington, 54 s.
- NIELSEN, T. A. S., HOVGESSEN, H. H. (2008): *Exploratory mapping of commuter flows in England and Wales*. Journal of Transport Geography, 16, s. 90–99.

- Python v2.6.5 documentation*. Python Software Foundation, 2010. Dostupné z: <http://docs.python.org/ftp/python/doc/2.6.5/python-2.6.5-docs-pdf-letter.zip> (cit. 15. 4. 2014).
- SciPy.org documentation*. Verze 0.14.0. The SciPy community, 2014. Dostupné z: <http://docs.scipy.org/doc/scipy/reference/> (cit. 15. 5. 2014).
- SMART, M. W. (1974): *Labour market areas: uses and definition*. Progress in Planning, 2, s. 239–353.
- SPURNÁ, P. (2008): *Geograficky vážená regrese: metoda analýzy prostorové nestacionarity geografických jevů*. Geografie - Sborník České geografické společnosti, 113, č. 2, s. 21–35.
- SÝKORA, L., MULÍČEK, O. (2009): *The micro-regional nature of functional urban areas (FUAs): lessons from the analysis of the Czech urban and regional system*. Urban Research & Practice, 2, č. 3, s. 287–307.
- TOBLER, W. (2004): *On the First Law of Geography: A Reply*. Annals of the Association of American Geographers, 94, č. 2, s. 304–310.
- VEČERKA, A. (2007): *Grafy a grafové algoritmy*. Univerzita Palackého, Přírodovědecká fakulta, Katedra informatiky, Olomouc, 106 s. Dostupné z: http://phoenix.inf.upol.cz/esf/ucebni/Grafy_a_grafove_algoritmy.pdf (cit. 13. 5. 2014).
- WATTS, M. (2009): *Rules versus Hierarchy: An Application of Fuzzy Set Theory to the Assessment of Spatial Grouping Techniques*. In: Kolehmainen, M., Toivanen, P., Beliczynski, B. (eds.): Adaptive and Natural Computing Algorithms. Springer Berlin Heidelberg, Berlin, s. 517–526.
- WATTS, M. (2013): *Assessing Different Spatial Grouping Algorithms: An Application to the Design of Australia's New Statistical Geography*. Spatial Economic Analysis, 8, č. 1, s. 92–112.
- WORBOYS, M. F., DUCKHAM, M. (2004): *GIS: A computing perspective*. 2. vydání. CRC Press, Boca Raton, 426 s.

Seznam příloh

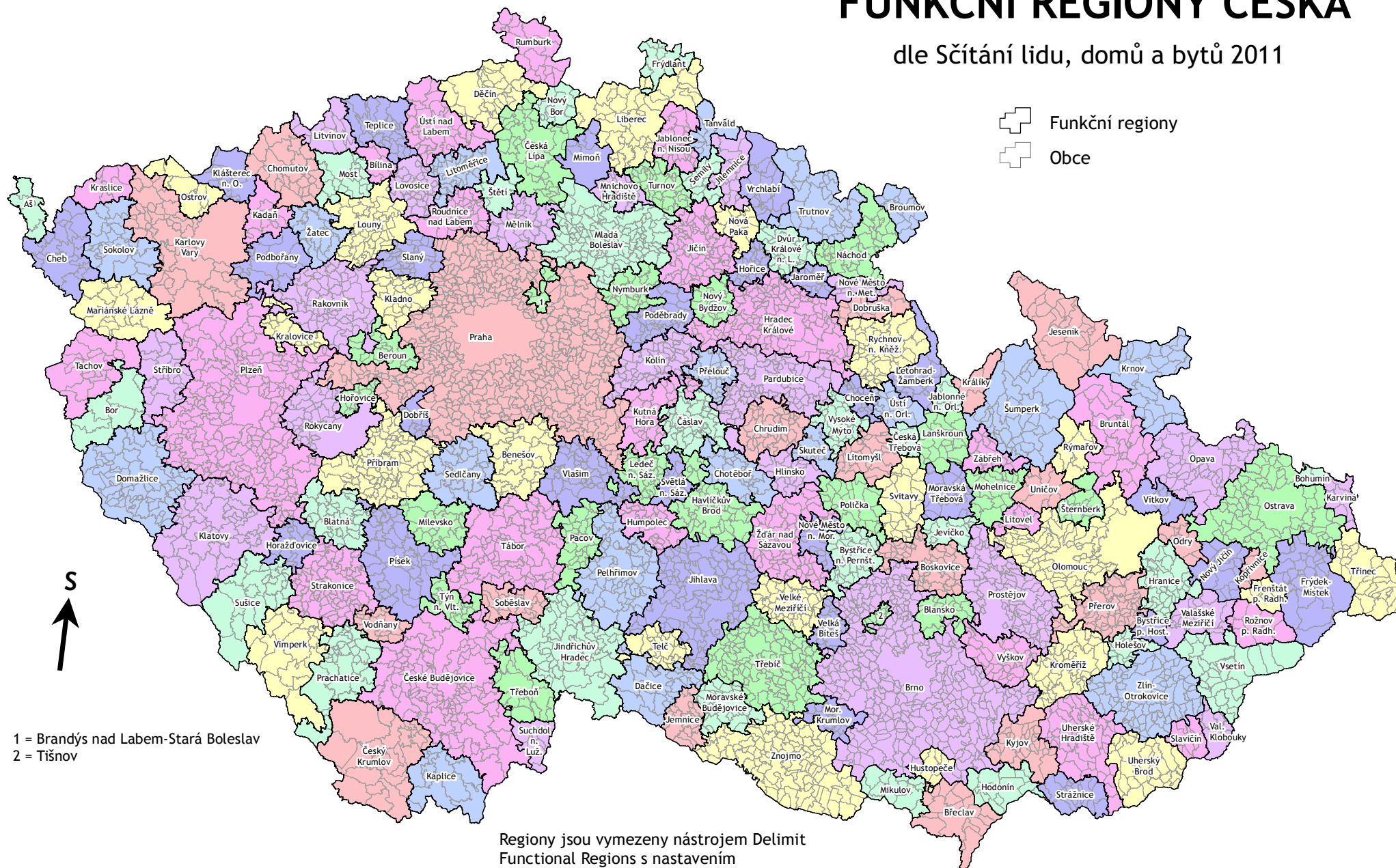
1	CD s toolboxem Interactions a testovacími daty (ČSÚ 2013)	v deskách
2	Struktura tříd nástroje Delimit Functional Regions	58
3	Regionalizace Česka 2011 – přehledová mapa	59
4	Regionalizace Česka 2011 – vizualizace příslušnosti zón k regionům	60
5	Regionalizace Česka 2011 – členské funkce obcí	61
6	Charakteristiky vymezených regionů – tabulka	62

Příloha 2: Diagram tříd nástroje Delimit Functional Regions.



FUNKČNÍ REGIONY ČESKA

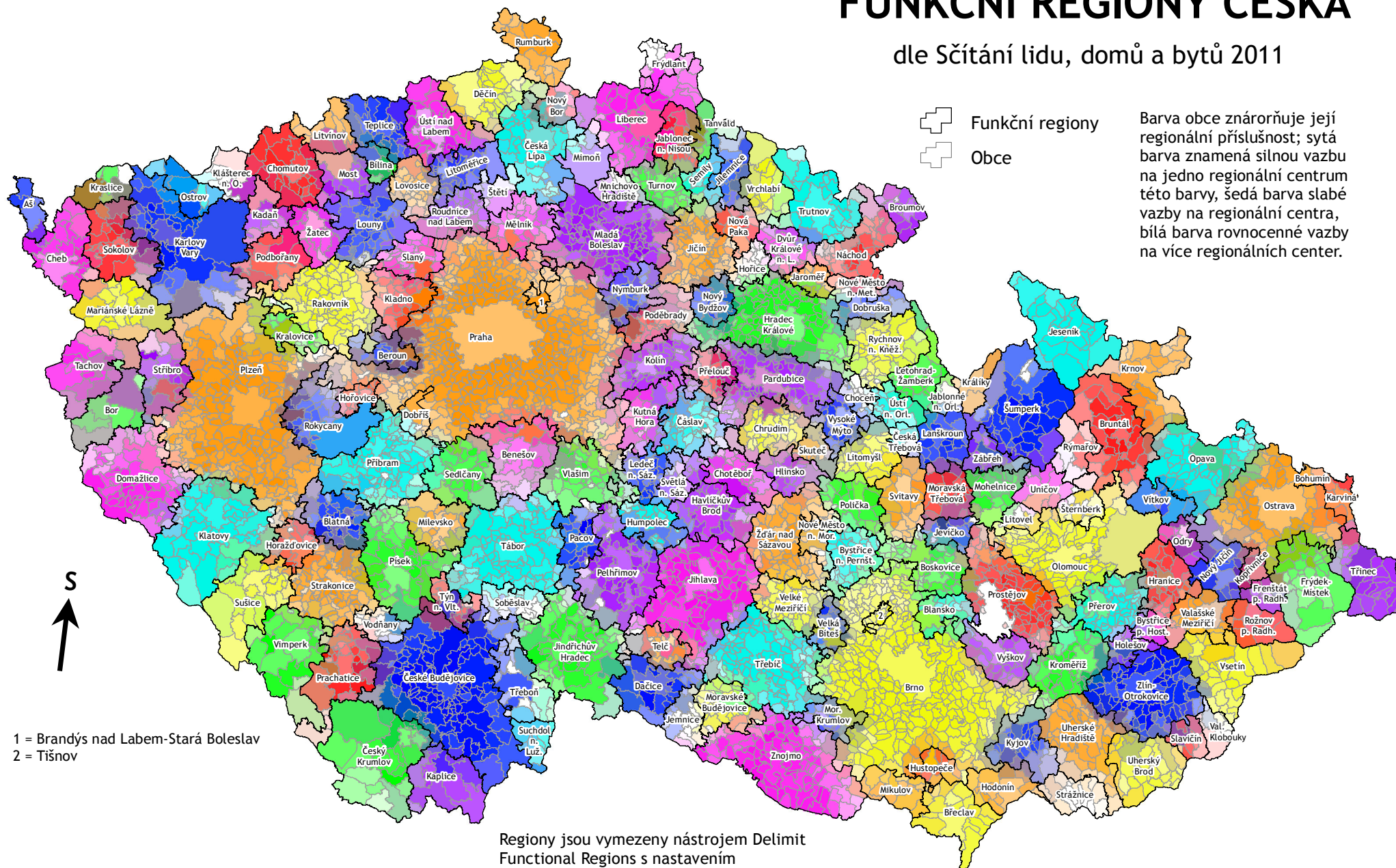
dle Sčítání lidu, domů a bytů 2011



1 = Brandýs nad Labem-Stará Boleslav
2 = Tišnov

FUNKČNÍ REGIONY ČESKA

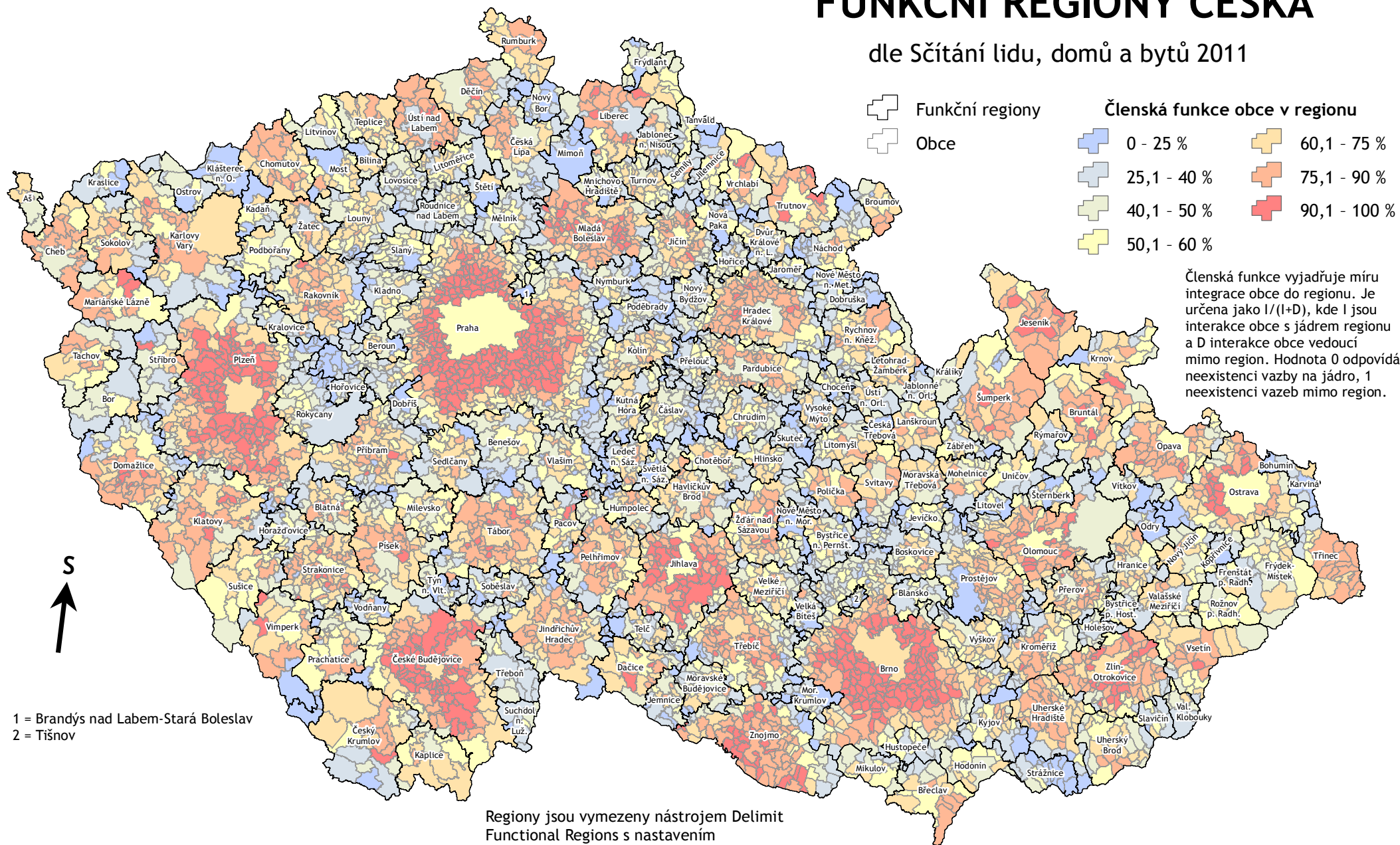
dle Sčítání lidu, domů a bytů 2011



Regiony jsou vymezeny nástrojem Delimit
 Functional Regions s nastavením
 $W_0 = 1000$ (min. členská váha zázemí),
 $k_{e0} = 0$ (nulová tolerance exkláv),
 $r_1 = 10\%$ (minimální přerazovací zisk),
 $RO_0 = 50\%$ (minimální překryv vícejaderných regionů).

FUNKČNÍ REGIONY ČESKA

dle Sčítání lidu, domů a bytů 2011



0 50 km
1 : 1 750 000

Regiony jsou vymezeny nástrojem Delimit
Functional Regions s nastavením
 $W_0 = 1000$ (min. členská váha zázemí),
 $k_{e0} = 0$ (nulová tolerance exkláv),
 $r_1 = 10 \%$ (minimální přerazovací zisk),
 $RO_0 = 50 \%$ (minimální překryv vícejaderných regionů).

Příloha 6: Charakteristiky vymezených regionů. I_R je Hamplova integrita regionu (viz část 2.6.1 na s. 19).

Název regionu	EMW	Obyvatel	Zón	I_R	Plocha (km^2)
Praha	1 143 670	1 791 964	473	3,52	4 290
Brno	507 485	700 969	325	3,76	3 016
Ostrava-Havířov	363 545	577 924	73	1,22	1 089
Plzeň	260 323	335 821	221	3,69	2 616
České Budějovice	151 204	201 778	127	3,56	1 892
Zlín-Otrokovice	121 761	172 090	82	2,44	826
Olomouc	116 115	185 527	74	2,16	1 177
Hradec Králové	98 375	171 831	116	1,50	1 015
Pardubice	76 826	156 711	96	1,00	752
Liberec	76 419	154 455	40	1,14	798
Opava	75 029	119 146	52	1,46	724
Mladá Boleslav-Kosmonosy	74 792	109 252	109	2,06	974
Jihlava	74 655	106 781	106	2,90	1 126
Uherské Hradiště	69 707	108 648	51	1,33	608
Třinec	68 265	104 477	26	1,31	453
Karlovy Vary	68 100	115 971	49	1,53	1 310
Znojmo	67 483	89 930	101	1,78	1 146
Teplice	66 957	109 958	28	0,92	355
Prostějov	60 193	101 497	86	1,09	809
Tábor-Sezimovo Ústí	57 243	82 713	81	1,67	1 030
Třebíč	56 062	84 556	103	1,64	942
Šumperk	55 130	78 808	42	1,98	918
Ústí nad Labem	54 102	123 788	28	1,13	478
Frýdek-Místek	52 597	108 845	32	0,74	650
Chomutov-Jirkov	50 154	82 183	25	0,70	486
Příbram	45 097	74 006	81	1,12	778
Přerov	43 791	85 599	63	0,90	428
Vsetín	42 478	65 570	31	1,36	643
Trutnov	41 107	65 061	31	1,34	602
Děčín	40 901	76 962	31	0,90	505
Kroměříž	40 242	71 490	45	1,02	477
Rumburk	39 194	54 368	18	2,58	355
Klatovy	37 827	55 782	53	1,60	1 096
Kladno	36 740	110 911	32	0,35	316
Sokolov	35 835	61 991	27	0,87	494
Břeclav	34 442	58 444	17	1,13	423
Náchod	33 193	64 954	39	0,96	429
Jeseník	33 186	40 657	24	3,64	719
Česká Lípa	32 612	68 523	38	0,79	618
Domažlice	32 460	47 742	66	1,80	893
Hodonín	32 317	60 631	15	0,95	270
Kolín	30 991	62 488	51	0,82	414
Jindřichův Hradec	30 660	48 514	61	1,26	967
Žďár nad Sázavou	29 060	49 861	65	1,26	614
Cheb	29 014	51 101	20	1,39	455
Uherský Brod	28 928	50 170	28	0,99	460
Rakovník	28 767	45 781	69	1,17	761
Strakonice	28 299	46 075	74	1,36	608
Písek	27 878	51 468	46	1,07	741
Rychnov nad Kněžnou	26 944	46 146	42	1,30	505
Jablonec nad Nisou	26 520	71 468	24	0,49	259

...

Název regionu	EMW	Obyvatel	Zón	I_R	Plocha (km^2)
Pelhřimov	25 888	42 698	65	1,32	752
Jičín	25 701	47 241	74	1,05	519
Bruntál	25 066	42 234	32	1,33	713
Nový Jičín	23 597	48 352	16	0,74	275
Louny	23 505	44 366	44	0,78	509
Havlíčkův Brod	22 907	45 984	48	0,83	515
Krnov	22 091	42 001	25	0,90	575
Vyškov	21 967	43 925	35	0,73	335
Kyjov	21 507	46 043	32	0,76	360
Chrudim	20 261	48 462	42	0,59	345
Valašské Meziříčí	19 574	44 547	20	0,85	297
Karviná	19 391	73 219	6	0,36	124
Český Krumlov	19 174	35 394	24	1,07	993
Boskovice	18 999	39 498	57	0,81	394
Litoměřice	18 833	45 843	30	0,54	309
Hranice	18 578	36 480	35	0,88	350
Mariánské Lázně	18 555	29 329	17	1,17	516
Rožnov pod Radhoštěm	17 722	36 288	10	0,71	247
Turnov	17 671	32 687	38	0,96	251
Vrchlabí	17 186	30 436	18	1,27	327
Rokycany	16 503	36 036	39	0,63	573
Svitavy	15 622	32 222	28	0,80	362
Most	15 594	76 365	18	0,34	261
Litvínov	15 514	39 573	11	0,60	236
Benešov	14 889	38 717	25	0,49	527
Velké Meziříčí	14 554	26 493	43	0,95	343
Lanškroun	14 322	23 571	23	1,36	295
Prachatice	14 004	26 336	33	0,91	565
Tachov	13 705	26 128	18	0,88	533
Čáslav	13 621	32 650	45	0,58	360
Blansko	13 243	37 307	20	0,42	207
Litomyšl	13 113	24 556	29	0,93	306
Sedlčany	13 079	22 705	22	0,99	455
Polička	13 012	21 909	25	1,12	316
Mohelnice	12 768	24 080	20	1,03	274
Letohrad-Žamberk	12 727	22 726	21	0,92	272
Kutná Hora	12 607	35 120	29	0,49	361
Mělník	12 143	34 535	26	0,45	339
Vlašim	12 035	25 618	48	0,67	490
Sušice	12 010	21 411	24	1,00	631
Roudnice nad Labem	11 630	29 872	31	0,51	286
Beroun-Králův Dvůr	11 266	34 300	17	0,21	217
Dvůr Králové nad Labem	11 234	26 608	24	0,58	233
Kopřivnice-Štramberk	11 187	37 651	6	0,31	85
Vimperk	10 991	18 793	24	1,13	688
Lovosice	10 321	26 647	30	0,51	246
Mikulov	10 123	19 309	16	0,84	235
Slaný	10 078	27 704	36	0,39	250
Chotěboř	10 061	22 029	28	0,65	315
Dačice	10 039	17 486	26	1,19	426
Žatec	9 972	26 169	15	0,64	260
Hlinsko	9 751	19 482	21	0,74	224
Bystřice nad Pernštejnem	9 688	19 236	34	0,78	308
Nymburk	9 637	28 950	25	0,41	255

...

Název regionu	EMW	Obyvatel	Zón	I_R	Plocha (km^2)
Moravská Třebová	9 579	19 041	18	0,84	263
Ústí nad Orlicí	9 553	25 492	16	0,60	179
Humpolec	8 874	18 423	27	0,86	255
Uničov	8 762	22 959	11	0,52	213
Jilemnice	8 676	20 167	18	0,64	249
Milevsko	8 528	17 118	24	0,73	336
Aš	8 513	17 635	5	0,87	144
Vysoké Mýto	8 505	22 203	32	0,63	230
Blatná	8 487	16 496	31	0,90	368
Poděbrady	8 376	29 322	32	0,31	321
Broumov	8 061	14 666	12	0,92	184
Podbořany	8 045	16 045	13	0,79	354
Kaplice	7 996	15 084	13	0,88	457
Zábřeh	7 823	22 992	18	0,42	177
Moravské Budějovice	7 702	16 482	30	0,73	287
Frenštát pod Radhoštěm	7 700	21 274	7	0,50	119
Přelouč	7 308	20 734	26	0,44	192
Semily	7 126	18 176	14	0,55	151
Nová Paka	7 083	19 477	13	0,48	176
Holešov	7 027	20 751	18	0,39	128
Nový Bydžov	7 023	16 334	18	0,60	185
Stříbro	6 891	16 405	20	0,51	396
Hořovice	6 811	17 887	20	0,44	99
Desná-Tanvald	6 296	17 023	7	0,32	148
Valašské Klobouky	6 042	15 536	10	0,53	141
Strážnice	5 972	23 028	17	0,29	241
Ostrov	5 971	24 324	8	0,31	238
Ledeč nad Sázavou	5 590	11 121	25	0,75	230
Pacov	5 586	12 060	27	0,63	297
Bystřice pod Hostýnem	5 563	13 595	9	0,37	97
Třeboň	5 431	16 375	17	0,40	374
Jaroměř	5 363	18 074	13	0,36	122
Dobruška	5 316	15 446	16	0,45	188
Česká Třebová	4 784	19 893	6	0,33	90
Hustopeče	4 735	12 797	9	0,58	109
Soběslav	4 704	13 226	22	0,43	224
Jevíčko	4 695	11 279	15	0,60	153
Nový Bor	4 692	20 296	11	0,24	156
Slavičín	4 687	13 206	10	0,44	139
Hořice	4 586	14 196	21	0,41	138
Telč	4 474	12 182	27	0,48	218
Rýmařov	4 285	11 553	8	0,53	226
Horažďovice	4 196	10 515	16	0,56	194
Litovel	4 109	14 958	8	0,34	113
Kraslice	4 039	11 491	6	0,41	201
Bohumín	4 029	27 058	2	0,19	56
Kadaň	3 850	22 594	10	0,26	218
Šternberk	3 844	17 888	13	0,23	178
Nové Město nad Metují	3 714	15 415	18	0,26	136
Vodňany	3 695	11 348	14	0,43	161
Nové Město na Moravě	3 651	14 055	13	0,31	149
Vítkov	3 605	10 914	7	0,40	194
Mnichovo Hradiště	3 586	12 913	17	0,33	150
Jemnice	3 550	8 392	21	0,51	207

...

Název regionu	EMW	Obyvatel	Zón	I_R	Plocha (km^2)
Králíky	3 543	8 462	4	0,54	146
Bílina-Ledvice	3 531	18 915	6	0,22	114
Choceň	3 444	12 833	17	0,30	109
Frýdlant	3 441	12 080	9	0,30	160
Skuteč	3 440	12 583	11	0,32	143
Moravský Krumlov	3 425	11 045	17	0,34	174
Odry	3 409	11 380	8	0,39	145
Bor	3 372	9 677	8	0,51	405
Mimoň	3 340	14 145	6	0,29	241
Klášterec nad Ohří	3 309	21 446	9	0,14	231
Týn nad Vltavou	3 152	11 413	12	0,26	187
Velká Bíteš	3 108	9 235	19	0,41	145
Světlá nad Sázavou	3 088	10 510	13	0,33	138
Suchdol nad Lužnicí	3 071	8 814	7	0,53	158
Štětí	3 065	12 680	7	0,30	150
Brandýs n. L.-St. Boleslav	2 932	21 283	8	0,15	75
Jablunné nad Orlicí	2 891	7 768	10	0,44	93
Kralovice	2 885	7 226	21	0,50	194
Dobříš	2 862	12 415	7	0,24	111
Tišnov	2 318	11 683	8	0,13	48